The Dissertation Committee for Nathan Kant McVaugh certifies that this is the approved version of the following dissertation

The Effect of Instant Messaging on Lecture Retention

Committee:

____

Daniel H. Robinson, Supervisor

____

Erika A. Patall

____

Diane L. Schallert

____

Marilla D. Svinicki

____

Georghios Veletsianos

The Effect of Instant Messaging on Lecture Retention

by

Nathan Kant McVaugh, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2012

Dedication

To Lydia, my family, and teachers everywhere.

Acknowledgements

It would have been impossible for me to complete this dissertation without my partner, Lydia J. Middendorf. Thank you for the help, understanding and patience, which you've shown throughout the years of this endeavor. I could not have done this without you, and now our investment has paid off. Thank you, my love.

I would also like to thank my parents and family for their support throughout this process. While my father did not survive to see my completion, he knew the outcome was just a matter of time. I'm proud to have answered the educational and service calling which our family continues to be dedicated to.

The unceasing efforts of the department of educational psychology were obviously essential. The faculty of the educational psychology department (both current and retired) continues to do an outstanding job of guiding aspiring researchers through the process of becoming the next generation of researchers. My special thanks to my dissertation committee – Dan Robinson, Erika Patall, Diane Schallert, Marilla Svinicki and George Veletsianos. Your work is reflected in this and scores of other dissertations including your own.

Equal thanks to the staff members who have contributed just as much to the process, with very special thanks to Virginia Stockwell, Julian Chapa, Rob Donald, James Keys, and all the staff at the Casteñeda Library Inter–library Loans department. As a fellow staff member at UT Austin, I want to thank you personally for everything you do to make the university run smoothly and for running interference for the grad students.

Many, many thanks for everything you all have done for me and the other education students. You are the unsung heroes of the university.

My deepest thanks to those who have gone before me and have sought to fit their understanding to the world, rather than the world to their understanding. To name but a few, Plato, Galileo Galilee, Nicholas Copernicus, David Hume, Charles Darwin, Karl Popper, and Richard Dawkins. In the words of R.E.M., you "had the gall to ask." Thank you for providing the shoulders to stand on and pointing towards the future.

Finally, my thanks to all the under–appreciated, ridiculed, disdained and disregarded public employees throughout the world. Thanks to those who strive to make a difference each day in the lives of strangers, and are rewarded with scorn for your efforts. The military, firefighters and police lack nothing in lip service, though public support falls short for you as well. So much the worse then for teachers, paramedics, social workers, defense lawyers, and the countless other professionals who have invested years of toil and personal wealth to pursue the dream of building a better society for all. Your work is appreciated, admired and essential – thanks for shouldering the burden of keeping the light of civilization burning.

Those wishing to support this process are urged to contribute to classroom science teachers throughout the US at this address http://www.donorschoose.org/.

The Effect of Instant Messaging on Lecture Retention

Nathan Kant McVaugh, Ph.D.

The University of Texas at Austin, 2012

Supervisor: Daniel H. Robinson

The impact of instant message interruptions via computer on immediate lecture retention for college students was examined. While watching a 24–minute video of a classroom lecture, students received various numbers of related–to–lecture ("Is consistent use of the eye contact method necessary for success?") versus not–related–to lecture ("Have you ever missed class because you couldn't find parking?") instant messages in addition to note taking vs. no note taking. Student self–rating for multitasking ability, typical and maximum instant messaging activity, and classroom computer use were also measured. Contrary to cognitive models of information processing that suggest instant messages will disrupt student retention of lecture information, no effects were found for number of interruptions, presence or absence of notes, or relatedness of interruption on lecture retention. Students' multitasking self–rating was negatively related to lecture retention. The implications of these results for classroom practice and future research are explored.

Table of Contents

List of Tables

List of Figures

1. Introduction

Problem Statement

The incorporation of new technologies into the higher education classroom has always engendered controversy. Partisans of emerging technologies suggest a revolution in education is imminent, with adoption of the latest technological tools helping to usher in a new era of learning. Although the widespread adoption of wireless internet technologies has produced remarkable tools to address research and communications problems, it has simultaneously engendered a new set of classroom issues. Instructors must now compete with a wide variety of electronic distractions for student attention in the classroom, raising concerns about the negative impact of these technologies on classroom learning. A chief concern is the well–established finding in cognitive psychology of severe limits on human information processing (Baddeley, 2007). Decades of experiments have shown how division of attention between simultaneous tasks results in declines in execution speed and accuracy for those tasks. Given the these cognitive constraints, it seems reasonable for instructors to expect the prohibition of laptops and cell phones in the classroom would be associated with better student performance, or alternately would preclude a decrease in student performance. This result would be consistent with some correlational and experimental studies, which have found laptop use to be associated with decreased academic performance (Fried, 2008; Hembrooke & Gay, 2003; Kirschner & Karpinski, 2010).

This experiment examines the effect of instant message interruption on student retention of lecture information. As the number of instant message interruptions increased, it was expected that students would be forced to devote more attention to the messages, leaving less attention for lecture comprehension and retention.

Were the issue solely one of divided attention, this would likely have been the case. However, classroom learning is a complex phenomenon involving the interaction of many factors besides limited attentional capacity. Along with varying the number of interruptions, the content of the messages was also varied, with half the subjects receiving messages related to the lecture content, while the other half received social messages unrelated to the lecture.

Further analysis and refinement of the experimental manipulations suggests additional factors, which should be included when considering the impact of instant messaging on lecture retention. Whereas the pedagogical benefits of note taking both with paper and on computer (Oyzon & Olmos, 2010; Williams & Eggert, 2003) are well documented, studies have not examined the impact of note taking in conjunction with instant messaging. This is especially relevant as students cite note taking as a primary use of computers in the classroom, along with referring to classroom materials such as the syllabus and course web site. Presumably the positive effects of note taking interact with other factors such as interruption by instant message. However, it is unclear how much of an interaction there is, and whether the combination of note taking and instant message interruption are positively or negatively associated with classroom learning outcomes.

Similarly, student arousal is known to affect learning, with too much or too little arousal having a negative impact, while a moderate level has a positive impact (Eysenck & Calvo, 1992; Eysenck, Derakshan, Santos, & Calvo, 2007). A large number of instant messages have the potential to distract students, but a few messages throughout the course of an otherwise uninspiring lecture may have a beneficial effect on student alertness and lecture retention. The relationship between instant message interruption and student alertness has yet to be addressed in any formal manner.

The research covered here examines a series of experiments, the results of which indicate concerns and expectations about the impact of technology interruptions in the classroom may be overblown. In this and other experiments, technology's impact on immediate recall was extremely variable. It appears that the impact of interruptions is moderated by a variety of individual and situational variables. As a result, blanket bans on technology in the classroom cannot guarantee positive learning outcomes, and in some cases may be detrimental. While additional factors could be considered, this analysis illustrates the complex issues raised by wireless technologies in the classroom. I now turn to the reasoning behind this study, and how it addresses technological interruptions, note taking, and student arousal in a higher–education classroom.

Purpose of the Study

This study examined how classroom technology, including instant messages and note taking, affected recall of a 24–min. video clip covering the Eye Contact Method for dealing with classroom disruption (Jones, 1978). As unanticipated results appeared the experimental procedure was adjusted to address additional issues. In total, three separate experiments examined the impact of factors such as the frequency of instant message interruption and the relatedness of the interruption to the lecture content (Pilot, Preliminary and Concluding study), and the presence or absence of a field for taking notes in the software the students used during the experiment (Pilot and Concluding study). Results from the Preliminary study indicated the role of student arousal should be considered, and this issue was specifically addressed in the Concluding study.

In addition to instant message manipulations, each experiment included instruments to measure factors of general interest. The Pilot study included measures for student interest in the content of lecture and for multitasking ability. The Preliminary study included a self–report measure for how students used computers in the classroom. The Concluding study included a measure of introversion/extroversion. Taken together the experiments measure the impact of computerized distractions such as instant messaging on immediate recall of a classroom lecture. Specifically, the experiments addressed the following questions.

Research Questions and Hypotheses

This study addresses seven principal questions and hypotheses:

1. Based on the principle that cognitive processing limits are more likely to be exceeded when cognitive load increases, it is hypothesized that larger numbers of messages will produce greater processing deficits, resulting in lower scores on the recall quiz. This hypothesis takes the form of the question, "Does the frequency of IM interruption have an impact on lecture recall?"

2. Based on the principle that information unrelated to a topic interferes with the cognitive processing, it is hypothesized that instant messages unrelated to the lecture will produce greater processing interference than related messages. This is expected to result in lower scores on the recall quiz for unrelated messages than related ones. This

hypothesis takes the form of the question, “Do IM questions relatedness to the lecture have an impact on lecture recall?”

3. Based on the principle that active processing of information through note taking aids in retention of that information, it is hypothesized that taking notes will result in higher scores on the recall quiz. This hypothesis takes the form of the question, “Does the opportunity to take notes affect lecture recall?”

4. Based on the principle that there is an optimum level of cognitive arousal and that external stimuli may move students towards or past that level, it is hypothesized that students with low cognitive arousal will show recall quiz benefits from interruptions, while those with high arousal will have impaired performance. This hypothesis takes the form of the question, “Do cognitive arousal traits affect lecture recall?”

5. While the above factors independently impact cognitive processing load, it is possible that interactions between these factors may result in one factor enhancing or attenuating the impact of another. Thus it is hypothesized that interaction effects beyond those of individual factors will appear as measured by scores on the recall quiz. This hypothesis takes the form of the question, “Are there any significant interactions between these factors?”

6. How students use computers in the classroom is an ongoing area of research, with a general consensus that most students rely on computers for note taking while a few use them for distraction. It is hypothesized that students in this experiment will report computer use consistent with this pattern, leading to the question, “How do students use computers in classroom situations?”

7. Based on the observation that people often have unrealistically high estimates of their abilities, it is hypothesized that students who estimate they have higher multitasking proficiency will have lower scores on the recall quiz than those who do not claim such proficiency. This hypothesis takes the form of the question, “How accurate is student self–assessment of multitasking ability?”

Significance of the Study

This study serves to extend our understanding of how communications technology affects student performance in the college classroom, but also serves as a reminder of the complexity of cognition in the classroom, requiring researchers and instructors to consider multiple factors when trying to anticipate how technology will impact learning.

This study incorporates several components in service of this goal, with applicability to actual undergraduate classrooms being foremost. A video recording of a typical college lecture presentation was used to test performance, rather than an artificial sorting or resource allocation task. This format approximates a classroom environment for assessing student attention and performance in the classroom, allowing factors such as student interest in the lecture topic and their desire for distraction to manifest in realistic ways.

Although the psychological literature has repeatedly demonstrated limits on attention and processing capacity, this does not necessarily imply that eliminating distraction in the classroom will lead to improved recognition or recall of lecture material. The first research question tests this idea by varying the number of instant message interruptions to see if there is a resulting change in immediate lecture recall.

The second question depends on the first – if there is an effect of interruption, is it necessarily a negative one? Timing the interruptions to appear at specific times in the lecture made it possible to send messages related to the lecture point being raised, or to send an unrelated message that interrupts the key point being made.

Turning to the third question, one of the most cited uses of classroom computers is for taking notes during class. The Preliminary study did not include note taking, as this was deemed to be irrelevant to the question of interruption. Given the apparent importance of arousal, note taking was added to the Concluding study to assess how this factor interacted with interruptions. If student recall is suffering due to lack of stimulation, then the process of taking notes should serve to provide additional stimulation. However, overstimulation would be expected to occur if notes are combined with a high number of interruptions, resulting in poorer recall. This experiment manipulated note taking across several interruption conditions to explore this factor.

The importance of a realistic classroom lecture was dramatically illustrated during the Primary experiment when one student fell asleep during the lecture. In his defense, he said the lecture was not sufficiently stimulating, and the lack of interruptions by instant messages was responsible for his slumber. In isolation, this would be unremarkable, but during this experiment several other students exhibited signs of boredom, a result not addressed by the processing capacity model. This result suggests educational researchers should consider student arousal in classrooms as a potential moderator in many situations, including the evaluation of the effect of technology such as instant messages in the classroom. The Concluding study included a measure of student cognitive arousal to address the fourth research question.

As just noted, factors such as arousal and instant message interruptions would be expected to interact in affecting lecture recall. Analyzing the experimental results for interaction effects among all three variables (note taking, message–relatedness and message–frequency) addressed the fifth question.

The last two questions were addressed by gathering self–report data from students about how they use computers in classrooms, and how well their scores match with their estimated ability to multi–task.

Finally, this study contributes to the literature describing student use and attitudes towards technology in the classroom. Self–report data about laptop use from the Laptop Effectiveness Scale (LES) are used with a different population from the original study (Lauricella & Kay, 2010). Questions about how students perceive their own technology use are included. Within the context of existing studies, this information deepens our understanding of how students perceive and approach technology.

The combination of a realistic stimulus and a balanced quantitative design is hoped to lead to further investigation of how interruptions interact with variables such as student arousal. The results of such studies will be of interest for both researchers and classroom educators, and are certain to become more relevant as technology continues to assert itself in the classroom.

2. Literature Review

Introduction

Much of higher education relies on the lecture method of instruction, with instructors reviewing and explaining material to their students, answering their questions and guiding them through the concepts and methods which comprise the focus of the class. Although this method has been used with a reasonable degree of success for hundreds of years, today's lecturers feel they must increasingly compete for student attention (Bugeja, 2006; 2007a; 2007b; 2008). To educate, many instructors feel they must vanquish electronic distractions in the classroom offered by games, instant messaging and social media sites (Foster, 2009; Mangan, 2001; Yamamoto, 2007; Young, 2006). For these educators, prohibiting the use of laptops and smart phones seems required to focus student attention on the difficult task of learning beyond a superficial level.

In contrast, others have promoted a conception of today's students as fundamentally and uniquely different from previous generations, and suggest current and future students have grown up in a world which has honed their cognitive and neurological development to allow simultaneous and effective use of multiple streams of information (Brown, 2000; Negroponte, 1995; Prensky, 2001a; 2001b; Veen & Ben Vrakking, 2006). This assessment is often echoed by mass media (Brooks, 2001; Meade, 2003; Wallis, 2006) and students, who describe themselves as bored and distracted when not working with multiple sources of information simultaneously (Bowman, Levine, Waite, & Gendron, 2010; Frand, 2000; Hammer et al., 2010; Watson & Strayer, 2010).

According to this view, instructors should abandon the lecture in favor of developing new teaching approaches catering to the technological mastery and need for stimulation which distinguishes the new generation of “digital natives.”

The competing utopian/dystopian narratives are well stocked with anecdotal and “common sense” claims, but quantitative research is relatively scarce in this area. How does mobile technology affect students in the classroom? Are today’s students able to effectively allocate their attention across multiple stimuli and process multiple streams of information? What classroom policies should instructors adopt to best educate their students, and what policies should they avoid? Most centrally, is technology a boon or burden in the classroom, and what factors are important in making this determination?

To answer these questions, we must consider the origin and role played by computer and communications technology in the lives of students, and the niche this technology occupies in the classroom and in their personal lives. In addition, cognitive, attentional and multitasking factors must be considered, with special attention to arguments about the abilities of “digital natives.” This background leads to a set of quantitative experiments to measure how these factors interact in a simulated classroom lecture to impact student learning and lecture retention.

Technology and Instant Messaging

In 1965, Intel co–founder Gordon E. Moore predicted that for at least the next decade, the density of transistors in integrated circuits would double every two years while the price stayed constant, resulting in a logarithmic increase in computer processing speed and memory capacity while costs remained unchanged (Moore, 1965).

Within a few years, this prediction had become enshrined within technological circles as "Moore's Law" and has proved to be surprisingly robust – in part because chip manufacturers have adopted research, development and marketing goals reflecting this "law" (Lente & Rip, 1998). Whatever the reason, the power of computing devices has increased logarithmically, to the point where laptop computers are capable of running hundreds of programs simultaneously and yet are so affordable they have achieved near universal adoption by higher education students (Caruso & Salaway, 2007; Diamanduros, Jenkins, & Downs, 2007; Smith, Caruso, & Kim, 2010; Smith, Salaway, & Caruso, 2009).

Taking advantage of this increase in computing power, computer networks have likewise shown phenomenal growth in capacity. Although their development lags behind increases in processing speed, most higher education institutions maintain computer networks capable of wirelessly connecting thousands (or tens of thousands at larger campuses) of computers to the internet (Mahometa, Eakin, & Smith, 2008). Simultaneously telecommunications companies such as AT&T and Verizon have moved to offer support for data (as opposed to voice) networking for computers and smart phones. Current public and private wireless networks can easily handle high definition video and audio transmissions, leaving plenty of bandwidth for more basic data formats such as text.

The result has been an explosion in computer–based text communications. Indeed, despite widely reported drops in literacy (Bradshaw & Nichols, 2004), estimates for individual textual consumption and production have increased in response to technology

(Griswold, McDonnell, & Wright, 2005; Mokhtari, Reichard, & Gardner, 2009), with text–intensive web sites leading the way.

Social networking software such as Facebook allows students to interact with over half a billion individuals and organizations online (Zuckerberg, 2010). Video streaming sites such as YouTube allow instant access to tens of millions of video clips, and while an exact count is impossible, in July of 2011 there were over 107 million active web sites (Netcraft, 2010).

On the educational side, resources such as Wikipedia currently offer around 3.7 million articles in English, with the technical articles being of surprisingly high currency and quality (Giles, 2005; Hu, Lim, Sun, Lauw, & Vuong, 2007). Search and indexing systems for peer–reviewed literature and technical information such as Web of Science, PsychInfo and Google Scholar allow academics access to a more complete and current collection of research than any campus library could house (Bar–Ilan, 2008; Jacso, 2005).

Finally, application programs such as spreadsheets, presentation software, word processors and statistical programs allow students to access and manage information in ways not possible just a few years ago. When used in combination with online resources, students and instructors have unprecedented ability to access, capture and share information in the classroom (Ryder, 2000; Wang, Shen, Novak, & Pan, 2008).

Following web sites, email and the more interactive technology of “instant” or text messaging contribute to increases in the amount of reading and writing done on a daily basis (Leu et al., 2007; Liu, 2005), with instant messaging being preferred to email by some college–age and younger users (Flanagin, 2005).

Instant messaging resembles email in allowing messages to be exchanged over the Internet between registered users, regardless of operating system or platform (computer or smart phone), with no “per message” charge for use. Unlike email, the messages are exchanged in a synchronous, real–time manner best suited to short, informal content. As such, it has become extremely popular for social interactions among today’s students (Battestini, Setlur, & Sohn, 2010; Lenhart, Rainie, & Lewis, 2001; Shiu & Lenhart, 2004). Free instant messaging should be distinguished from text messaging where short messages are sent and received, usually over cell phones, and may incur a per message charge.

Student Population

Today’s student body is well–equipped to utilize this technological infrastructure. While some campuses have mandated laptop purchases for students (Zucker, 2004) or provided them in cooperation with computer manufacturers (Hembrooke & Gay, 2003), these kinds of requirements may not be necessary to ensure students are “wired.” The most recent annual ECAR survey of technology on higher education campuses (Smith et al., 2010) indicates computer ownership has remained steady from 2004 to 2010, with over 98% of students reporting computer ownership. During this time, the proportion of laptops has increased by about 25%, while desktops have declined by 25%. In 2010, 89% of students reported owning either a laptop or notebook computer, while 46% reported owning a desktop computer (39% reported owning more than one computer). Of these students, slightly more than half (56%) described their computers as one year old or less, and only 17% reported their computers were four years old or older. Despite the ongoing economic malaise, those students who can still afford college are willing and able to acquire current computer technology such as laptops, cell and smart phones, and are bringing it into the classroom.

The adoption of newer technology is reflected in changes to cell phone ownership. While smartphones accounted for only 21% of wireless subscribers at the end of 2009, they are expected to account for the majority of subscribers by the third quarter of 2011 (Entner, 2010), and the number of smart phones is expected to exceed the number of people on the planet by the end of 2012 (Cisco, 2012). At present, academic use of this technology in classrooms lags behind availability, with less than 20% of instructors incorporating it into lectures (Smith et al., 2010). In contrast, over 70% of undergraduate students use instant messaging multiple times per week, and over 90% use text messaging on a daily basis, despite per–message charges.

Looking more closely at student use of electronic technology, a study by Flanagin (2005) found subject pool undergraduates undergoing a shift in technology use. Self–report measures correlated IM use with decreases in the use of other media, specifically drops of 71% for landline telephone use and 38% for email, though it should be noted 47% indicated no change in use of email. A nationwide survey conducted for the Pew Institute (Shiu & Lenhart, 2004) revealed similar technology use patterns by 18 to 27 year olds. Specifically, 62% reported IM use, and 20% reporting daily use. As a proportion of IM users, this age group accounts for 31% of all IM use. They also provide usage figures for IM, with 47% of IM users reporting 15 or fewer minutes at a time, 26%

reporting between 15 minutes and one hour of use at a time, and 22% reporting IM interactions lasing over an hour.

More recent reports indicate daily instant messaging has declined recently, though this may be due to changes in how “instant messaging” is conceptualized. As Smith (2010) reports:

> The dominance of texting and SMS use appears to be paralleled by a decline in instant messaging. Since 2007, the percentage of students who use instant messaging daily has dropped from 48.0% to 23.7%. However, this reported decline may be an issue of nomenclature. Formerly dominant instant messaging programs, such as Windows Live Messenger and ICQ, are likely losing ground to the functionality in other communications programs such as SNSs and VoIP—what the younger generation calls “chat” today comes from Facebook, Google, or Skype—and it may be that “instant messaging” as a stand–alone tool does not resonate with these students. (p. 62)

Clearly, students are using the technology on a regular basis. The ECAR studies also reported in 2009 over half the students surveyed (5%) currently own an Internet capable handheld device such as a “smart phone” or iPod Touch. In addition, another 12% intended to purchase such a device within the next year. This intention proved accurate, with the 2010 survey respondents indicating 63% owned these devices, and 11% expecting to purchase one in the next year. While these figures suggest large Internet use for these devices, at the same time more than one–third (35%) of device owners reported they did not utilize these features (Smith et al., 2009; 2010), though this seems likely to change as technology is developed, refined and popularized.

Technology in Student Lives

While computers are beneficial (if not essential) in higher education, their integration into classroom lecture and learning is often uneven. Technology such as clickers is becoming more common in classrooms, but in most classrooms instructors use computers primarily as overhead projectors (Young, 2006), rather than to promote interaction between the instructor and students. Outside of “clickers,” instructors seldom use computers to respond to students in the classroom (Caldwell, 2007; Connor, 2009; Guerrero, 2010; Watkins & Sabella, 2008).

On the student side, the most common use of computers is to take class notes (Lauricella & Kay, 2010, McCreary, 2009). There is a great deal of empirical support for note taking in the classroom. Taking notes using computers has been consistently tied with superior learning performance as measured by classroom exams (Baker & Lombardi, 1985; Katayama, Shambaugh, & Doctor, 2010; Oyzon & Olmos, 2010; Quade, 1996; Williams & Eggert, 2003).

The use of computers to facilitate note taking is the single most cited use of laptops in the classroom. In one study (Lauricella & Kay, 2010), student self–reports indicated 40% used their laptop for note taking for 76% to 100% of class time, with an additional 34% of students reporting such use from 51% to 75% of the time. Another study (Borbone, 2009) found 66% of students reported taking notes as the primary classroom use for their computers. In a survey measuring student laptop use and reactions to bans of laptops in the law classroom (McCreary, 2009), 80% of students reported using laptops in class on a regular basis, with 96% of laptop users reporting they used them to take notes in the classroom. Among students who had been in a class with laptops banned, 72% of laptop users felt the notes they took by hand were not as good as when they used a laptop.

Regarding the type of notes students write, concerns have been raised about the format technology makes possible. For example, typing is generally faster than writing by hand. Noting this, some educators have suggested the "transcription" style of notes possible on computers results in inferior learning outcomes (Read, 2011; Yamamoto, 2007).

In fact, research indicates note taking styles vary from sparse outlines to near "transcriptions" (Hadwin, Kirby, & Woodhouse, 1999), but generally students who take extensive notes on the computer do as well in class as students who do not use computers. In one experiment, students engaged in a 2 x 3 (pencil and paper, online notepad; free style, verbatim or paraphrase) with a no–note control condition experiment where they took notes over four consecutive lectures. Verbatim note takers (both online and pencil and paper) scored significantly higher than free style and paraphrase groups on subsequent recall tests (Quade, 1996). In another study, student notes from an introductory psychology lecture three weeks prior to a scheduled exam were collected and compared with their scores on the exam. Note completeness was significantly related to performance on the exam, with notes relating to quiz questions being a sufficient but not necessary condition for correctly answering test questions (Baker & Lombardi, 1985).

Despite its potential, technology has proved to be a mixed blessing in the classroom. Proponents can point to note taking and technological programs to improve or expand the classroom experience (Vogel, Kennedy, & Kwok, 2007; Wang et al., 2008), but technology has made it equally easy for students to access course–related and non–related materials online. Rather than take notes or reference course material, bored students may now interact with friends through instant messaging, exchanging email, or interact with the technology by playing video games or surfing the web (Leach, Lewin, & Pearson, 2007; Young, 2006). Many students do report boredom and lack of interest in the classroom. In his study McCreary (2009) found 45% of law students reported they used the computer because they were bored in class. At a technical college, Hammer et al. (2010) found 97% of students reported using laptops for non–classroom activities when bored, and 74% used cell phones.

Expanding on these numbers, the 2009 ECAR study (Smith et al., 2009) found almost one–third of respondents (32%) either agreed or strongly agreed that in class they regularly use their handheld device for "non–course activities." Another study (Lauricella & Kay, 2010) reported 56% of students surveyed spend up to half of class time sending and receiving instant messages, and over 70% of students spend up to half of class time in non–academic activities. This occurred despite the fact that instructors in their study made considerable efforts to engage the students (Kay, personal communication). Fried (2008) found several significant negative correlations between laptop use and various classroom measures, including final grades, understanding of the lecture, and attention to the lecture.

Studies such as these have been heavily promoted in the popular press (Brooks, 2001; Levy, 2006; McWilliams, 2005) and in the *Chronicle of Higher Education*, which has consistently drawn attention to the darker side of technology in education (Bugeja, 2006; 2007a; 2007b; 2008; Foster, 2009; Levy & Nardick, 2011; Mangan, 2001; Olsen, 2002; Young, 2004; 2006). Students have contributed to this view with comments such as the following from a 17 year old being typical, though hard to verify: "I multi–task every single second I am online. At this very moment, I am watching TV, checking my email every two minutes, reading a newsgroup about who shot JFK, burning some music to a CD and writing this message" (Lenhart et al., 2001).

In response, some instructors have opted to ban laptops and handheld devices from the classroom (Foster, 2009; Fried, 2008; Mangan, 2001), preempting both the costs and benefits of these technologies. Indeed, for some instructors removing technology from the classroom has been adopted as a kind of moral imperative, which should be embraced by responsible instructors everywhere (Maxwell, 2007; Yamamoto, 2007; YouTube, 2010). This attitude (and the popular accounts which support it) has been described as an academic "moral panic" (Bennett, Maton, & Kervin, 2008; Wallis, 2010), based upon the almost evangelical approach that many laptop opponents have embraced. Indeed, statements from earlier studies indicate that the same concerns and counter–arguments were present decades ago, and have returned with each passing generation (Anderson, Levin, & Lorch, 1977; Morgan, 1980). As Hornik (1981) noted "Did the replacement of comic book reading by television viewing and the consequent reduction in reading practice slow the learning of reading skills?" A conceptual overview of this issue can be seen in Figure 1.

While incarnations of these two polarized camps have dominated the discussion of technology in classrooms for several decades, our understanding of the actual impacts of technology has lagged behind. Most recently researchers have begun to examine the evidence for both a "net generation" of learners and for the quantifiable effects of classroom technology on education. With regard to the "net generation," the initial hype about students who think and learn in fundamentally different ways from previous generations (Prensky, 2001a; 2001b; Tapscott, 1998; Veen & Ben Vrakking, 2006) has given way to reviews which point to both a lack of evidence for unique learning preferences, and considerable variability of learning styles within the "net generation" itself (Bennett et al., 2008; Bullen, Morgan, & Qayyum, 2010; Helsper & Eynon, 2010). While students spend more time with technology than previous generations, this does not appear to have resulted in unique abilities to divide attention, or "multitask," as generational comparisons indicate (Carrier, Cheever, Rosen, Benitez, & Chang, 2009). Empirical studies of multitasking and attention switching abilities show the cognitive and neurological limitations of a generation, which has grown up in an environment of networked technology, do not differ from previous generations (Bennett et al., 2008; Bullen et al., 2010). Instead, the same limits on attention and memory observed decades ago (Miller, 1956) still appear to be in effect (Cowan, 2001). To understand how technology can impact student learning in the classroom, I will next briefly consider current models of memory, cognition and attention.

Memory Architecture

Many of the earliest models of human cognition were developed in response to a central set of experimental observations, specifically those which emerged from serial recall tasks where a subject must observe, retain and then recall a series of stimuli. Measurement of disrupted cognitive processing has produced several decades of research, which has tested and refined models of human information processing while continuing to produce new findings. Initial experiments examined retention for sequences of letters and digits (both of which remain popular choices), but the stimuli are limited only by a researcher's ingenuity. The spatial location of items, sequences of colors, musical tones, and many other variables have also been tested. All show patterns of interference from specific stimuli or tasks, illuminating an underlying cognitive architecture to support such processing.

Baddeley and his colleagues proposed one of the most popular and productive models of working memory (Baddeley, 1986; 2000; 2007). This model has proved so productive that even its critics refer to it as the "standard model" of working memory (Nairne, 2002). The model initially emerged from the pioneering work of Atkinson and Shiffrin (1968) and postulated a system composed of a sensory register, a limited short–term memory store, and an unlimited long–term memory store. Baddeley's "multicomponent model" (Baddeley 2007, pp. 7–10) initially postulated a three part system consisting of a "central executive" to perform general cognitive and attentional processing; a "phonological loop" to store acoustic and phonemic information; and a “visuospatial sketchpad” for storing and integrating visual and spatial information (Baddeley, 2007; Baddeley & Hitch, 1974).

Baddeley conceived of the phonological and visuospatial systems as operating independently. For example, in reading the visuospatial component would transfer the visual information to the central executive, which would then translate the visual information into its phonological components. Similarly auditory input, hearing the command “look up” for example, would require phonological information to be transferred into the central executive, which would then rely upon the visuospatial system to comply with the request. Note that in both the case of reading and of understanding language, we rely on largely automated processes with require very little attention or cognitive effort.

Because these two systems operate independently, overloading one system would be expected to have little impact on the other. Experimental tests of this hypothesis have produced a range of significant, and often contradictory, interference effects (Baddeley, 2007; Baddeley & Hitch, 1974; Craik, Govoni, Naveh–Benjamin, & Anderson, 1996; Naveh–Benjamin, Guez, & Marom, 2003; Naveh–Benjamin, Kilb, & Fisher, 2006).

The earlier work of Baddeley and Hitch (1974) was developed in an environment where earlier research had been conducted in the context of the effects of workplace music on productivity. An early example of a multitasking study was conducted with 26 college–age (18 to 23 year old) low–skill assembly line workers at a skateboard factory (Newman, Hunt, & Rhodes, 1966). Over a five–week period, control periods of no music were alternated with 4 varieties of music, balancing for day of the week effects with a

Latin–square design. Although workers overwhelmingly enjoyed the music, their productivity and error rate were unchanged. These results reflected those of McGehee and Gardner (1949), who found music made no difference in the productivity of highly skilled rug–setters, though it was preferred to silence. A review of the literature by Uhrbrock (1961) found that while most employees (90% to 99%) enjoyed workplace music, beneficial results in productivity were less clear–cut. Specifically, some studies indicated that quality of work could decrease when music was added. Contrary to Newman et al. (1966), some studies reviewed by Uhrbrock found that low–skilled employees engaged in simple, repetitive tasks did increase productivity. Consistent with McGehee and Gardner (1949), the review found experienced workers involved in complex tasks did not increase productivity in response to music.

In part to address findings of this nature, Baddeley proposed dividing the central executive between a long–term memory system and an attentional component called the "episodic buffer" (Baddeley, 2000; 2007; Baddeley & Logie, 1999), which acts as a central switchboard connecting the central executive, phonological and visuospatial components. By directing attention to information active in the other components, the episodic buffer binds the separate elements together to from chunks as first described by Miller (1956).

Multitasking

Further research in this area (Baddeley & Logie, 1999; see Baddeley 2007 for a review, especially pp. 117–138) found subjects with a heavy concurrent load on both the phonological and visuospatial systems performed well on recall and processing tasks, but only so long as they were not required to switch attention repeatedly between these tasks. When subjects were required to alternate attention repeatedly, or “multitask,” some measures of performance can be significantly poorer (Alderman, 1996; Armstrong & Greenberg, 1990; Baddeley, Sala, Papagono, & Spinnler, 1997; Hartman, Pickering, & Wilson, 1992; Logie, Cocchini, Delia Sala, & Baddeley, 2004). Unfortunately, extending laboratory research findings on multitasking into the classroom can be problematic.

Much of the older research is designed to test specific models of attention. The result is methodologically rigorous, but of little use to instructors. For example, Altmann and Trafton (2002) used a Towers of Hanoi (also known as Towers of London) puzzle as the task, and latency in task resumption as a measure of cognitive load. While the results were significant, it is unclear how to translate them into classroom practice. Similarly, a quasi–experiment by Ophir, Nass and Wagner (2009) correlated heavy self–reported multitasking in subjects to increased distractibility and lower accuracy in the context of recalling the orientation of a red rectangle on a computer display and applying logical rules to letter combinations. Additional studies have used a complex, strategy game based on resource allocation as the experimental task (Altmann & Trafton, 2004; Trafton, Altmann, Brock, & Mintz, 2003).

Other research is inapplicable because it lacks any quantitative measure or model. For example, Bowman, Levine, Waite & Gendron (2010) state “Students often believe they can listen to music, watch TV and/or communicate with friends online while doing school work without any detriment to performance.” However no data are offered, and it is unclear what is meant by “believe,” or how “performance” is defined and measured. A similar criticism applies to Watson and Strayer's (2010) assertion, "Over the years, we have encountered a great many people who adamantly claim that they are not impaired when they use a cell phone while driving." As people generally do a poor job estimating their own cognitive abilities (Kruger & Dunning, 1999), application of a more accurate assessment method is desirable.

Kinzie, Whitaker and Hofer (2005) attempted to combine classroom learning with instant messaging, having students instant message a learning partner for three weeks during an instructional technology course. During the course, students were assigned to discuss (via IM) potential uses and drawbacks of the instructional technology being covered in class that day. Post–class analysis of the messages found about one half of the messages dealt with the lecture topic, though there were significant differences between pairs of subjects in the amount of time spent in on–topic discussion. Feedback from participants revealed 68% of students would not recommend IM during class, reportedly viewing such discussions as best separated from the regular lecture. They also report students do acknowledge a trade–off between multitasking and task performance, again based on informal observations rather than controlled studies. Excluding artificial or impressionistic cases, there have been mixed and conflicting experimental results when looking for multitasking deficits in learning. Considering prior literature reviews (Freeburne & Fleischer, 1952), this has been the case for the last 60 years.

Mixed results were reported by Pool et al. (2000), who performed two 2 X 3 mixed block design experiments on the effect of native–language (Dutch) soap opera, English MTV or no television program on high and low difficulty homework assignments. The first experiment found no impact of background television on combined high and low difficulty homework scores; though students in the soap opera condition did score significantly lower on the difficult homework assignment. No differential effects for time to complete homework were found.

The second experiment replicated the first, but added an additional factor of differential instructions regarding the television and a test of television clip recall. Half the subjects were instructed to ignore the television, while the other half were told, "When you do your homework in front of the television, you usually want to know what is happening in the television program. Here, it will be the same. You should concentrate on the assignments, but you may keep an eye on what is occurring on the television screen at the same time." There was no significant difference in homework responses for either television condition versus control. There was a significant effect for time, with students watching the soap opera taking significantly longer (22%) than the other two groups. There was no significant main effect of television instructions on clip recall, though there was an interaction effect with MTV viewers showing better clip recall than the ignore television group.

Another experiment (Armstrong & Greenberg, 1990) also used television as a distractor across seven cognitive processing tasks. Significant results were only found for the more difficult operations of reading comprehension, spatial problem solving, and cognitive flexibility. No significant effects were found for the other four, less demanding conditions. This is consistent with performance deficits emerging under higher cognitive load conditions where a change in modality is accompanied by attention switching (Baddeley, 2007).

Fox and colleagues (Fox, Rosen, & Crawford, 2009) examined reading times and comprehension in psychology undergraduates who either did or did not instant message while reading. No difference was found in recall or recognition scores between the groups, though students in the IM condition did take significantly longer both to read passages and to answer recall questions afterwards.

A similar study (Bowman et al., 2010) compared reading time and quiz performance on students who did not instant message or who instant messaged before or after reading a passage with those who used instant messaging while reading the passage. Again, recall scores did not significantly differ, with the authors noting "No differences in test performance (number correct) were observed for those who IMed before, those who IMed during, and those who did not IM" (p. 930). And once again reading times for the IM while reading condition were significantly higher than the other conditions.

A study by Hembrooke and Gay (2003) has gained significant attention for linking laptop use in the classroom to lecture recall. In the study, two classes were given the same lecture, and then tested on recall. Students in the control condition kept their laptops closed during the lecture, while students in the second condition were allowed to use their laptops as they saw fit. Following the lecture students were given a 20–item quiz on the lecture, which included ten multiple–choice (recognition) questions and ten short–answer (recall) questions. Two months later a follow–up study using the same students reversed the group assignment, and a second ten–item lecture quiz was given.

The study showed a significant negative effect of laptop use on recognition score for the first ten–item lecture quiz, and for the second ten–item lecture quiz with the reversed groups, though recall score was unaffected. So far as quiz scores go, it would appear normal laptop use in a lecture can have a negative impact on recognition scores, though not on recall.

Hembrooke and Gay further analyzed how students were using their computers during the lecture, dichotomizing students into “ontaskers” if they spent more than 50% of their online time viewing course–related content, and “offtaskers” if they spent less than 50% of their online time this way. Comparison of on versus off–task browsing for “ontaskers” and “offtaskers” found significant differences in the times the two groups spent on related and unrelated web pages, indicating the division reflected real differences in how these groups were using the web.

While “offtaskers” switched between course–related and unrelated web sites, they spent about the same number of minutes on each kind of page. In contrast, “ontaskers” spent three times as long on unrelated pages as they did on related pages. When quiz scores were dichotomized across “ontaskers” and “offtaskers,” the on–task students total quiz scores were significantly lower than those of the off–task students. Additional studies will be needed to replicate and specify the factors responsible for these results. At present, it appears individual differences within a treatment condition play a greater role in moderating the effect of technology use on lecture retention than previously suspected.

Together, the Pool et al. (2000), Bowman et al. (2010), and Fox, Rosen, and Crawford (2009) studies found no evidence that the kind of multitasking that students commonly engage in while studying (television and instant messaging) has a quantifiable impact on performance. The Armstrong and Greenberg (1990) study only found results for more difficult conditions, but the tasks themselves were not typical of classroom conditions. Only Hembrooke and Gay (2003) found an impact for laptop use, and this was for recognition (though not on recall) scores for a lecture quiz.

Classical and recent cognitive research on multitasking has shown that while we can switch attention between distinct tasks, and with practice we may even learn to do so with some efficiency (Dzubak, 2008; Felton, 2007), doing so is not without cost. Significant deterioration of performance, understanding and reaction time as the complexity and number of laboratory tasks increases has been found (Marois & Ivanoff, 2005; Meyer & Kieras, 1997; Pashler, Johnston, & Ruthruff, 2001), though results have been mixed and non–significant results are common. Recent classroom, homework and reading studies have shown increases in task duration, but have produced mixed results for measures of recall and recognition, indicating that the classroom and homework environments differ significantly from the purer findings of the research lab.

These results appear to be mirrored in the classroom. Given the large number of students that have laptops, course grades should have systematically dropped over the last two decades as computers became increasingly common in the classroom. Instead, grades have increased over this period (Rojstaczer, 2010). Similarly Hembrooke and Gay's (2003) results showed poor exam performance associated with computer use was not tied to final course grades in their upper–level communications class. As mentioned previously, all students in this class were provided with laptops for classroom use during the semester. So with the experimental exception of a single class period, students used the laptops in class throughout the semester. The universal presence of computers would be predicted to decrease grades, but the class earned a B+ average final grade.

These results may be understood by returning to Baddleley's (Baddeley, 2007; Baddeley & Logie, 1999) observation that significant performance deficits tend to be associated with studies in which students had demanding multitasking requirements across modalities. Significant results for quiz performance are often (but not always) associated with such high–difficulty tasks with modality switching. Significant results for execution speed are usually (but not always) associated with lower difficulty tasks or a single modality. Educational lectures require extended phonological and occasional visuospatial processing, and presumably deal with relatively difficult subject matter. Laptop computers are typically used to access written information, which imposes a low load on visual processing due to the automatic nature of reading. As a result, students in the classroom are usually not mixing modalities, and so a significant detriment would not be expected to take place without some additional factor becoming involved. One possible candidate is a psychological factor essential to classroom learning – student engagement.

Engagement and Cognitive Arousal

As anyone who has been in a classroom or attended a conference can attest, attention levels can vary widely. A PowerPoint lecture taking place in a warm, dimly lit room just after lunch is guaranteed to result in several examples of diminished attention, even in the absence of personal technological distractions. This lack of cognitive arousal or engagement results in decreased attentional capacity, reducing the resources available for learning.

Pioneering work in this area was performed early in the last century (Yerkes & Dodson, 1908), and found the level of physiological arousal was tied to task performance, including learning. This relationship was described as an inverted U shape, with extremely low or high levels of arousal hindering performance, and moderate levels of arousal enhancing performance on a given task. Also of note is that the inverted U shape predicts equivalent levels of performance when low and high arousal levels are equidistant from the apex of the curve (Broadbent, 1965; Malmo, 1957). Conversely, when dealing with tasks of varying difficulty, an inverse relationship is expected. Performance of difficult tasks requiring more concentration was best with lower levels of arousal, while easier tasks requiring less concentration and more motivation benefitted from higher levels of arousal. These results proved so robust it is now commonly referred to as the "Yerkes–Dodson Law." Unfortunately, the "law" fails to provide a mechanism to explain the relationship, as well and fails to specify how "arousal" and "difficulty" can be quantified (Eysenck, 1981).

There have been several attempts to address these shortcomings, with updated models describing arousal as being composed of two separate but interacting memory systems (Broadbent, 1971; Easterbrook, 1959; Kahneman, 1973; and Thayer, 1967; 1970; 1978). Hans Jürgen Eysenck created one of the most popular models, which claimed that individuals varied in average activity of the ascending reticular activating system (Eysenck, 1967; Eysenck, 1976). Individuals with higher levels of activation were considered to be near or above the peak of the Yerkes–Dobson curve (Anderson & Revelle, 1982). As a result any increase in cognitive activity would result in poorer performance on tasks, leading these individuals to avoid additional environmental stimulation. Conversely, individuals with lower levels of activation were at or below the peak of the curve, and could benefit performance by seeking to increase their level of environmental stimulation. Eysenck referred to the higher activity stimulation avoiders as "introverts", and the lower activity stimulation seekers as "extraverts." Unfortunately these labels are often conceptualized as being exclusively social in nature. Extraverts are viewed as more friendly, outgoing and talkative than introverts, but this is not commonly viewed as a reflection of an individual's typical level of cortical arousal. To emphasize the neurological aspect of the construct, "extraversion" terms will be replaced with the more descriptive "cognitive arousal."

The cognitive arousal construct has proved a fruitful area for experimental research, with hundreds of studies conducted in the decades since its introduction. The construct itself is a dimension of Eysenck's Personality Inventory (EPI), and is measured using a using a version of the Eysenck Personality Questionnaire (EPQ). There are several versions of the test available, with educational researchers often using Eysenck and Eysenck's (1994) 48–item Eysenck Personality Questionnaire – Revised Short Scale (EPQR–S). Different versions of the EPQ have been used in studies seeking correlations from cognitive arousal to student behaviors and outcomes. For example, Campbell and Hawley (1982) predicted there would be a correlation between a student's cognitive arousal score and where they chose to study in a campus library. They suspected students with lower cognitive arousal (extraverts) would seek to increase arousal level by frequenting the busier and noisier floor of the university library. In contrast students with higher cognitive arousal (introverts) would seek out the two quieter floors. In November of 1978 and 1979, students were sampled on each floor about five preferences associated with cognitive arousal. In one–tailed tests of significance, four of the five measures reached values of $p < .01$, with the remaining measures reaching $p < .05$.

A crossed study examining the interactions of impulsivity (a dimension of Eysenck's introversion/extraversion measure) and caffeine on student speed and accuracy in detecting two types of proofreading errors found several significant interactions (Anderson & Revelle, 1982). Within the experiment, low impulsive subjects performed better with a placebo than with caffeine, while high impulsive subjects performed the same or better with caffeine than without. In other words, increasing stimulation had differential effects depending on individual cognitive arousal.

On the physiological side, level of arousal has been quantified through measures such as EEG, Galvanic Skin Response, pupillary dilation, and heart rate. For reviews of the physiological literature see Gale (1973), Geen (1976) and Stelmack & Wilson (1982), as well as more recent neurological studies (Cohen, Young, Baek, Kessler, & Ranganath, 2005; Johnson et al., 1999). On the cognitive side, level of introversion has been linked to a variety of mental tasks including vigilance, attentional control and differences in short and longer–term recall (Eysenck, 1976; Eysenck et al., 2007; Furnham & Bradley, 1997; Howarth & Eysenck, 1968).

Combined Effects of Cognitive Arousal and Multitasking on Learning

As mentioned previously, the rapid expansion of technology in the classroom has been accompanied by experimental investigations into the impact of technology on student achievement, and an increasing focus on the role of cognitive arousal as a factor in student multitasking with technology. Students with low cognitive arousal (extraverts), would be expected to seek additional stimulation via multitasking, pushing them closer to their optimal level of arousal on the Yerkes–Dodson curve. As a result, having additional stimulation in the form of laptop distractions may improve some students' learning outcomes. Conversely, for student with high cognitive arousal (introverts) additional stimulation would be expected to push them farther away from their optimal level of arousal, resulting in decreased student learning outcomes.

This is not a new idea. Mackworth (1969) suggested that the introduction of irrelevant stimulation into a monotonous task could improve the subject productivity. Hill (1975) extended this observation to include the construct of cognitive arousal. To do so, a group of undergraduates were divided on the basis of extraversion scores. Both groups were then given the reasonably monotonous task of using a single push pin from each of three different sources to fill in 150 identical targets on a paper grid. Students could choose how to fill in the target, with the prediction that those with low cognitive arousal would show greater variety in how they filled in the target than those with higher arousal. A significant difference was observed, with low cognitive arousal subjects (extraverts) showing an average of 53.1 changes in response pattern compared with the 29.5 changes in response pattern by those with higher cognitive arousal (introverts).

In reviewing the literature on student multitasking, Dzubak wrote, “multitasking for most students is far less boring than uninterrupted focus on study. Young people and perhaps even many adults have become uncomfortable with silence and a lack of near constant environmental stimulation. Multitasking breaks the boredom” and continued, “It might be more pleasurable to multitask while studying but that does not correlate with more learning taking place” (Dzubak, 2008, p. 11). These comments assumed that all people respond to the demands of multitasking in the same way. In light of Eysenck’s arousal construct this comment would seem to be applicable to those with low cognitive arousal, but not to those with higher cognitive arousal. However it should be noted that multitasking while studying does not correlate with less learning taking place either. Unfortunately, a review of the literature shows that inclusion of cognitive arousal measures in experiments does not consistently result in significant results.

A series of three experiments tested the effects of background TV and radio on various tasks. Furnham et al. (Furnham, Gunter, & Peterson, 1994) ran a 2 X 2 between subjects design to look for a differential effect of background television on reading comprehension for those who scored low and high in levels of cognitive arousal. The comprehension tests consisted of two short (400 word) passages followed by multiple–choice questions. Each subject completed one question with television and one without, with presentation order balanced by group and test question. Post–test questions about how distracting students found the television were significant, with high cognitive arousal subjects saying it was more distracting than those with low cognitive arousal. They also found that regardless of arousal level, both groups did better on the tests with television off than not. What is more important, they found a significant interaction between television and personality, with high cognitive arousal subjects scoring significantly lower with television on, but no significant difference when there was no distraction. Note that in this case, significant results were found when students were juggling television (visuospatial) and reading (phonological) information sources, requiring multitasking between modalities. In this case, cognitive arousal appears to explain the differential results by predicting how and why specific groups of students will experience significant interference from multitasking.

Furnham and Bradley (1997) extended this design with a second experiment to see whether student cognitive arousal interacted with either silence or pop music to affect scores in two different recall tests. They used a 2 X 2 between subjects design to examine how student level of cognitive arousal interacted with background radio music to affect immediate and delayed recall of text and pictures. For the immediate reading recall condition, they found no arousal effect or interaction effect with music, but did find a main effect for presence or absence of music. The picture recall test occurred after a six–minute distractor task. This time there were no main effects for cognitive arousal level or the presence of music, but there was an interaction effect, with high cognitive arousal subjects significantly outperforming low cognitive arousal ones in the music condition. A significant correlation was found between subject level of cognitive arousal (as quantified by the EPQ scale) and how distracting students rated background music to be. Note that in this experiment the first (non–significant) interaction occurred with high load (music and text) only on the phonological loop, while the second (significant) interaction occurred with loads (music and picture) on both the phonological and visuospatial loop.

Furnham and Allass (1999) again extended this design, adding a third test of visuospatial reasoning to the previously used immediate and delayed visual recall tests and a comprehension test. The three tests were crossed with either silence, simple, or complex music as an audio distraction. They found a clear crossover effect for auditory distraction with visual recall, with high cognitive arousal subjects doing best in silence and worst with complex music, while low arousal subjects did worst with silence and best with complex music for both immediate recall and delayed recall. Combining auditory distraction with the visuospatial reasoning test also showed significant crossover deficits. However auditory distraction with reading comprehension tasks did not reach significance. High arousal subjects (introverts) rated complex (but not simple) music as being more distracting than low arousal subjects (extraverts).

A recent study (Gonder, 2010) looked for a connection between cognitive arousal, interruption via instant messaging, and retention of lecture content. Students watched a 50–minute classroom lecture on a computer screen with 0, 10 or 20 “text messages” appearing. Messages appeared at regular intervals – every 2.5 or 5 minutes for the 10 and 20–question conditions respectively. Students wrote down their responses to half of the messages on a sheet of paper, and all students had the option to take notes on a separate sheet of paper. Following the lecture students were quizzed on lecture content. For the purposes of analysis, the students were split along the dimension of high and low cognitive arousal as measured with the Eysenck Personality Questionnaire – Revised

Short Scale (Eysenck & Eysenck, 1994). No difference in recall quiz scores across interruption conditions was found. However, it should be noted that quiz scores for low arousal subjects in the 20–interruption condition were (non–significantly) higher than of high arousal subjects. As with Furnham and Bradley (1997), a single (verbal) memory system was used for both the lecture and recall tests. As a result, no switching between phonological and visuospatial systems was necessary.

Returning to Hembrooke and Gay (2003), cognitive arousal was not a variable in their model, and was not measured directly. However, if short but frequent web use by the "offtaskers" was the result of lower arousal students seeking stimulation, this would explain the lack of impact that this kind of browsing has on quiz scores. Conversely, the less frequent buy longer browsing of "ontaskers" would be consistent with higher cognitive arousal. Regarding processing, we again see significant results when students must juggle between lecture (phonological) and computer (visuospatial) tasks.

There was no difference due to cognitive arousal in the phonological distraction/ phonological problem (PP) condition for Furnham et al. (1994), Furnham & Allass (1999) or Gonder (2010). However, high cognitive arousal subjects did do worse than low arousal ones in Furnham and Bradley (1997). In all cases where there was a visuospatial task and an auditory task or distraction, high cognitive arousal subjects scored significantly lower than low arousal ones. Hembrooke & Gay (2003) did not examine level of cortical activity via the EPQ. However, their results and description of students are consistent with the lower scoring students having higher cognitive arousal.

Conclusion

Barring sudden social or economic changes, students who can afford to attend college will continue to have the latest technology, and universities will continue to embrace wireless and other technologies as academic necessities. Technology will continue to become smaller, cheaper, faster and more popular, practically guaranteeing a presence in the classroom. Do educators need educational policy to limit or ban technology use in the classroom to teach effectively? While this may be the case with some instructors (Maxwell, 2007; Yamamoto, 2007), is this an effective strategy for every instructor and class?

Given the limits on human cognition, it makes sense that laptops would compete for student attention, thereby limiting student learning. Yet the inclusion of multitasking with cognitive learning research has produced inconsistent and even contradictory results. This study will address some of these concerns by looking for a systematic impact of simulated instant message interruptions on short lecture recall. To do so, both the frequency and relevance of the interruptions will be varied, along with the student's opportunity to take notes. In addition student classroom computer use and multitasking ability will be surveyed.

3. Methodology

Overview

The current paper is based on a series of experiments stretching over three consecutive semesters. Conducted during the fall of 2010, the first experiment is referred to as the Pilot study. The second experiment was conducted during the spring of 2011 and is referred to as the Preliminary study. The third experiment was conducted during the fall of 2011, and is referred to as the Concluding study. When not described as part of a specific study, the methodology sections apply to all three studies.

Recruitment and Student Well–Being

Because the target population for this study is undergraduate students, students were recruited from the Department of Educational Psychology subject pool, and received course credit for participation. Additionally, the second study used students recruited from an educational psychology class who received extra credit for participation. Several steps were taken to ensure students provided informed consent throughout the research. All IRB protocols were followed, with emphasis on informed consent, freedom to decline participation, and risks and benefits of participation.

The experimental procedure utilized materials (computer and lecture) familiar to the students, and no undue stress was observed during the research. Students were aware of the purpose of the research – to investigate whether interruptions had any effect on recall of a lecture. This experimental approach was preferable to deception for two reasons. First, any resulting findings would be more realistic. Students may engage in technology use despite (or in some cases because of) the knowledge that doing so will distract from the lecture (Hammer et al., 2010; S. D. Smith et al., 2009; 2010). In addition, awareness of the purpose of the experiment addresses the criticism of students being able to easily guess the true nature of the experiment.

Students were further informed that the purpose of the experiment was to improve education by allowing instructors to develop technology policies based on research rather than anecdotal information. Contact information for the researcher and the IRB was provided both before and after the experiment, should students feel an ethical violation had occurred.

Hardware and Software

Most research was conducted in a computer lab that contained 30 Macintosh computers with Intel Core 2 Duo processors running version 10.6.8 (also known as Snow Leopard) of the Macintosh operating system. However, when scheduling conflicts occurred, a second smaller lab with 24 computers was used. Each lab included an instructor's computer with connection to a large screen for projection from the computer. This machine was utilized during the experiment for demonstrating the use of the software. Each student computer was equipped with a set of headphones, allowing students to listen to the lectures without being distracted by audio from other computers in the room.

The software consisted of a video file, a runtime environment, and a custom software program. The video file was a 24–min., 21–sec. excerpt of a lecture on classroom management techniques for dealing with disruptive students. The software for the experiment used the name of the software file to determine which experimental condition the software would enact. For example, in the Concluding study software with the name b3Stimulus would include the notes field and six instant message interruptions unrelated to the lecture content. In contrast, software with a file name of a5Stimulus would not display the note field, and the six instant messages would be related to the lecture content.

The software was developed in Revolution 4.5 from RunRev ltd. (http://www.runrev.com). This software is a descendant of earlier systems such as HyperCard, and allows rapid design and deployment of custom software without the complexity of more sophisticated programming languages such as C++.

Random Assignment

A unique method of randomizing student assignment to conditions was used in these experiments. As previously mentioned, the software for each experiment was stored on a flash drive. Because the flash drives were mass–produced, it was possible to create a sample of drives that were visually indistinguishable from each other. Before each experiment an opaque bag was loaded with sufficient drives to supply the students who had signed up for each experimental session. For example, in the Pilot study there were five experimental conditions. An experimental group of 18 students would therefore need four flash drives from each of the five conditions, giving 20 drives in total, with two left over.

A second advantage of using flash drives was the ability to use restricted random assignment to force unequal sample sizes (Shadish, Cook, & Campbell, 2002). In the previous example, at least two flash drives remained after the experimental session. If some students failed to attend, then additional flash drives would be left over. For the next experimental session, these drives were placed in the opaque bag along with enough flash drives to handle the students registered for the session. Doing so increased the likelihood students would end up in a condition lacking in participants, while ensuring each student would have a non–zero chance of being in any experimental condition (Shadish et al. 2002). By the end of the experiment, cell sizes had balanced out.

Because the proportion of students in some conditions differed between sessions, it is important to test for bias resulting from differences in subjects attending experimental sessions. For example, less motivated or less organized students often put off participation until the last minute, with the resulting bias applied unevenly in these later trials. This possibility was examined by comparing mean scores across experimental sessions in the results and analysis section.

Lab Preparation

Approximately half an hour before each session began, I arrived to set up the room for the experiment. This started by loading a short presentation of the experiment onto the instructor computer. Next, USB headphones were attached to the computers the students would be using for that experimental session. Each computer with headphones also received a flash drive.

As previously mentioned, flash drives were selected randomly without replacement from an opaque bag and attached to each machine used during the experiment. The instructor and students were blind as to which condition a given computer was assigned. When students arrived at the lab, they were allowed to choose any unoccupied machine. Thus while their choice of computers was not truly random, the condition of the computer was.

At this point the overhead projector was turned on and the presentation loaded to the first slide for the subject pool briefing. The slide contained the contact information for the researcher.

Procedural Overview

Briefing

Upon arrival at the computer lab, students were added to an attendance roster so they would receive subject pool credit for study participation. They were then told to sit at any computer that had headphones attached to it. Students arriving early were also told that the experiment would start about five minutes later than posted to allow for tardiness, that they should log into the Macintosh operating system, and that they were free to surf the web or use their cell phones while waiting for the experiment to start.

Once it looked like no more students would arrive, students were thanked for attending and the briefing was begun. Students were told about the purpose of the experiment, and were reminded of their IRB rights and options. After this, they were told they would be using a computer program to watch a classroom lecture, and would have to respond to instant messages when they appeared. They were explicitly told, “Depending on which computer you choose to sit at, you may receive a lot of messages, just a few, or none at all. If you do get instant messages, you need to respond to them. You do not have to write an essay – just show that you read the message and are responding to it.”

Following the recorded classroom lecture, they were told there would be a short multiple–choice quiz on the lecture content, along with demographic items. Students in the preliminary and concluding studies were also told that they would have a short survey item to complete as well (the LES scale and a subset of the EPQ–BV, both of which are described in the next section). Finally the students were reminded that the classroom lecture would last 24 minutes, and that answering the questions would take about ten to fifteen minutes, leaving plenty of time for students to get to their next class.

Next, students were shown Figure 2 so they could see what the software would look like, and the various parts of the interface were explained. Students were told they were using stereo headsets because they were all listening to the same lecture, and that it would get very noisy and make it hard to follow the lecture without them. Students were also shown which keys on the keyboard could be used to raise or lower the volume to a comfortable listening level.

Finally, risks, benefits, and experimental credit were reviewed to ensure informed consent had been obtained. Students were reminded that their responses were anonymous and that no confidential information would be asked of them. A call for any questions before beginning never received any follow up. At this point they were asked to turn off cell phones, and were shown how to activate their software.

Starting the Experiment

Using the large screen projection of the instructor's computer as a reference, students were shown how to locate the flash drive on their desktop, and double click it to see the contents, which consisted of three software files on the drive. Once it was confirmed that everyone had opened the window, students were shown how to double click on the experimental software file "with a letter, a number and the word 'Stimulus.rev' in it." Once they opened the program they were free to start the lecture.

At this point I turned off the overhead projector, did a quick check of the lab to ensure all computers were working properly, and addressed any issues with software or hardware that had emerged. If any students arrived too late for the orienting presentation, they were directed to the instructor's machine, where the previous information was again presented. They were then directed to choose a computer, and follow the procedure as described.

Experimental Interface

The software consisted of a window with a video clip and a button for starting the video. Just as it is impossible to pause or rewind a live lecture, no controls were included which would allow the students to pause or rewind the video clip. In addition, two windows could be displayed alongside the video window. If students were in a condition where notes were allowed, the "Notes" window was visible and could be used. If students did take notes, the notes were saved for further analysis.

The "Instant Message" window allowed students to receive and enter short messages, and was set to interrupt students by flashing and displaying a text message at specific times during the video clip. For each interruption, the words "TA message:" followed by the text of the interruption appeared in the message pane of the instant message window. The entire message window was alternately hidden and revealed five times, with a 1/10 of a second delay between each change in visibility. The rapid cycling of visibility effectively caused the instant message window to "flash" on and off to provide a visual distraction. To respond to the question, the student would type their answer in the instant message window and press the enter key. The words "Student response:" and the text of their answer would appear below the computer–generated question. An image of a typical lecture screen is shown in Figure 2.

A key experimental variable was the use of two types of interruptions. All interruptions were short questions of the kind that often appear in instant message exchanges. There were two types of questions, those related to the content of the lecture and those unrelated to the content of the lecture. The unrelated questions were intended to be social in nature, and related to local university and city events, popular television programs, planned semester break destinations and so forth. To create the related questions, the lecture was analyzed for the mention of concepts necessary to answer the quiz questions correctly. Once all these times were identified in the lecture, a question was created that reflected the content at that point of the lecture. Because important concepts were being introduced in the lecture at this point, instant messages would be expected to be especially disruptive (if unrelated) or especially helpful (if related) to scores on the lecture quiz (discussed shortly).

During the actual experiment, the software used this list of time–codes to randomly select the times when messages would appear for that subject, along with the appropriate message for that time. The random selection of times meant that subjects in identical conditions would not experience identical interruptions, merely an identical number of interruptions. Appendix A contains the complete list of related and unrelated questions.

The Experiment and the Conclusion

Immediately following the conclusion of the lecture, the students were taken to the first screen of the quiz, following which they answered any additional survey instruments, and finally to a set of demographic items. Students had to answer all questions on each screen before proceeding to the next one. There was no option to return to previous answers and change them. Once they had completed the demographic items, they arrived at the last screen (Figure 3), which informed them that the experiment was done and they were free to leave.

Lab Wrap–Up

As they finished, students were thanked for their participation, and any additional questions or comments were addressed. All computers left logged in were restarted; wiping any incidental data the students might have placed on the computer before the start of the experiment. Flash drives were collected for later data transfer, and headphones were collected from the lab. The instructor’s computer was logged out, and the lab was ready for the next reservation.

Measures

The measures described below addresses the seven principle questions and hypotheses:

1. Based on the principle that cognitive processing limits are more likely to be exceeded when cognitive load increases, it is hypothesized that larger numbers of

messages will produce greater processing deficits, resulting in lower scores on the recall quiz. This hypothesis takes the form of the question, “Does the frequency of IM interruption have an impact on lecture recall?”

2. Based on the principle that information unrelated to a topic interferes with the cognitive processing, it is hypothesized that instant messages unrelated to the lecture will produce greater processing interference than related messages. This is expected to result in lower scores on the recall quiz for unrelated messages than related ones. This hypothesis takes the form of the question, “Do IM questions relatedness to the lecture have an impact on lecture recall?”

3. Based on the principle that active processing of information through note taking aids in retention of that information, it is hypothesized that taking notes will result in higher scores on the recall quiz. This hypothesis takes the form of the question, “Does the opportunity to take notes affect lecture recall?”

4. Based on the principle that there is an optimum level of cognitive arousal and that external stimuli may move students towards or past that level, it is hypothesized that students with low cognitive arousal will show recall quiz benefits from interruptions, while those with high arousal will have impaired performance. This hypothesis takes the form of the question, “Do cognitive arousal traits affect lecture recall?”

5. While the above factors independently impact cognitive processing load, it is possible that interactions between these factors may result in one factor enhancing or attenuating the impact of another. Thus it is hypothesized that interaction effects beyond those of individual factors will appear as measured by scores on the recall quiz. This

hypothesis takes the form of the question, “Are there any significant interactions between these factors?”

6. How students use computers in the classroom is an ongoing area of research, with a general consensus that most students rely on computers for note taking while a few use them for distraction. It is hypothesized that students in this experiment will report computer use consistent with this pattern, leading to the question, “How do students use computers in classroom situations?”

7. Based on the observation that people often have unrealistically high estimates of their abilities, it is hypothesized that students who estimate they have higher multitasking proficiency will have lower scores on the recall quiz than those who do not claim such proficiency. This hypothesis takes the form of the question, “How accurate is student self–assessment of multitasking ability?”

Lecture Quiz

The Lecture Quiz was used to test recall of material covered in the video lecture, providing a way to measure the impact of message frequency, message relatedness and note taking on recall (experimental questions 1, 2 and 3). It consisted of ten multiple–choice items that appeared on the computer screen immediately after the end of video lecture. The initial four questions were presented, along with a continue button at the bottom of the screen. As students progressed through the questions a progress bar let them track how much of the quiz they had completed (see Figure 4).

Once the students had completed the first four items, they were taken to the next four. Once the students had completed the next four items, they were immediately taken to the last two questions of the quiz. There was no option to change previous answers. The quiz items are listed in Appendix B. The questions had all been written by the instructor whose lecture was used in the video, and were of the same form used in exams for his class.

Laptop Effectiveness Scale (LES)

The Preliminary study added a slightly modified version of a recently developed classroom laptop use scale (Lauricella & Kay, 2010). The scale is a self–report measure of academic and non–academic use of computers in the classroom. The original scale included two items regarding movie viewing, to which 89% of respondents said they never watched movies during class. However, the instrument did not include items for non–academic web browsing. The version used in this study replaced the movie items with questions regarding web browsing, and is designed to examine how students use computers in classrooms (experimental question 6) This scale is included in Appendix C.

As of this writing, there have been no publications that use this scale outside of the University of Ontario Institute of Technology. As a result, this study provides an opportunity to compare results from a different population to those gathered so far (Kay & Lauricella, 2011a; 2011b; Lauricella & Kay, 2010), including a study comparing LES scores on unstructured versus structured computer use (Kay & Lauricella, 2011c). The level of detail in the instrument provides the kind of fine–grained quantitative data about student technology use that is in short supply for educators. A sample screen from the Preliminary study is shown in Figure 5.

Eysenck Personality Questionnaire – Brief Version (EPQ–BV)

Eysenck and Eysenck's (1994) 48–item Eysenck Personality Questionnaire – Revised Short Scale (EPQR–S) and its predecessors have been used to measure personality factors including cognitive arousal for decades, and have a rich experimental literature. However, the complete question battery may provide more information than experimenters want. If just cognitive arousal is of interest, then Sato's (2005) Eysenck Personality Questionnaire – Brief Version (EPQ–BV) may be used. The EPQ–BV consists of 12 cognitive arousal (extraversion) items and 12 neuroticism items, all scored on a five–point Likert scale rather than the original yes/no format. Validation of the EPQ–BV (n=290) indicated coefficient alphas for cognitive arousal (extraversion) of .92, and test–retest reliability was .92 (Sato, 2005). Between the EPQR–S and the EPQ–BV the cognitive arousal (extraversion) scale correlated at .89 ($p < .001$). Reviews of the EPQ scale's lineage may be found in Francis, Brown, and Philipchalk (1992), Sato (2005) and Vingoe (1966; 1968). A validation using students (n = 685) in the UK, US, Canada and Australia compared the EPQ and EPQR–S, and found the forms offered equivalent psychometric performance (Francis et al., 1992).

To investigate the impact of cognitive arousal on recall (experimental question 4), the Concluding study used a slightly modified sub–scale of the EPQ–BV (Sato, 2005) to test for the construct of extraversion. Specifically, the scale was returned to a yes/no option, and the neuroticism portion was left out. The psychometric effects of these changes were discussed prior to the start of the study, but the modified instrument itself was not validated separately. The version used in the study can be seen in Figure 6.

Demographic and Miscellaneous Items

All three studies asked for the same demographic items. These items were age, sex, year in school (freshman through graduate student), and the college they were majoring in.

The Preliminary and Concluding studies also included questions about the lecture the students had just watched, as well as questions about their personal messaging habits. Whereas the Preliminary study item asked about how *similar* the lecture was to standard classroom lectures, the Concluding study item asked how *interesting* the lecture was. In addition students were questioned about their personal instant messaging habits in class and estimated their own ability to multitask (experimental questions 6 and 7). All items are listed in Appendix D.

1. Pilot Study

The Pilot study took place during the fall semester of 2010, and was intended to test the software and analysis methods that would be used in the primary studies. Students were also asked to rate their interest in the lecture topic and to rate their personal multitasking ability. A limitation of the pilot version of the software was its reliance on a survey web site to gather and process the student responses rather than process the data itself. As a result, it was not possible to keep copies of the notes or IM message responses that students provided.

Design

This study took place during the fall semester of 2010, and used a 2×3 factorial crossed design including a control. The dimensions were instant message contents (related versus unrelated) across interruption frequency (0, 18 or 36 messages) as the between–subjects factors. Note taking was available for all conditions. Note that because the control group received no messages, the instant message contents dimension (related versus unrelated) was collapsed for this group. The different combinations of treatments (message relatedness and frequency) addressed experimental question 5 regarding interactions between factors.

Participants

Of the students assigned from the subject pool to participate in this study, 188 elected to participate in exchange for receiving experimental credit. Of these, 22 cases were excluded from analysis when the data failed to record properly during the experiment, and no usable data were obtained. The remaining sample consisted of 101 female and 50 male students, with 5 opting not to respond. Age ranged from 17 to 25, with an average of 20.74 and a standard deviation of 1.36. There were 13 freshman, 18 sophomores, 23 juniors, 98 seniors and 4 graduate students. Data on area of study were not collected.

Procedure

Students watched the lecture and responded to instant messages when they appeared. Those who wished to do so took notes on the lecture. Immediately at the end of the lecture, students were taken to the first screen of the lecture quiz (see Figure 4), previously described in the Measures section.

Students worked through the three–page quiz, and concluded by answering demographic items (Figure 7). When finished, they were informed that their participation was complete and they were free to go (Figure 3).

2. Preliminary Study

This study took place during the spring semester of 2011, and incorporated several changes from the pilot. Specifically, the highest number of interruptions was reduced from 36 to a more realistic 18. In addition, the note taking field was removed, as it was felt to be irrelevant to the issue of interruption.

Design

Because the study was aimed at the effects of instant messages rather than note taking, the ability for students to take notes was removed from the Preliminary experiment. As a result, the Preliminary study used a 2×3 factorial crossed design including a control. The dimensions were instant message contents (related versus unrelated) and interruption frequency (0, 6 or 18 messages) as the between–subjects factors, with note taking not allowed in any condition. Note that because the control group received no messages, the instant message contents dimension is collapsed for this group. The different combinations of treatments (message relatedness and frequency) addressed experimental question 5 regarding interactions between factors.

Participants

Of the students assigned from the subject pool to participate in this study, 158 elected to participate in exchange for receiving experimental credit. In addition, 18 students from an educational psychology class also volunteered for the experiment and participated for extra class credit. Of these, two cases were excluded from analysis. In the first case, the data failed to record properly during the experiment, and no usable data were obtained. A second student slept through most of the experiment, and his results were excluded. The sample consisted of 94 female and 80 male students. Ages ranged from 18 to 47, with an average of 21.43 and a standard deviation of 3.67. The median and modal ages were both 21. There were 27 freshman, 33 sophomores, 29 juniors, 82 seniors and 3 graduate students. Students came from an assortment of majors, with 3 Architecture, 23 Business, 23 Communication, 28 Education, 2 Engineering, 7 Fine Arts, 1 Geosciences, 44 Liberal Arts, 28 Natural Sciences, 1 Nursing, 1 Social Work, and 13 Undeclared.

Procedure

As with the Pilot study, students watched the lecture and responded to instant messages when they appeared. As the note taking condition had been eliminated, students were unable to take notes. Immediately at the end of the lecture, students were taken to the first screen of the lecture quiz (see Figure 4), previously described in the Measures section.

Students worked through the three–page quiz, and were then taken to the three–screen LES scale (Figure 5). After finishing this, they concluded by answering demographic items (Figure 7). When finished they were informed that their participation was complete and they were free to go (Figure 3).

3. Concluding Study

The Concluding study took place during the fall semester of 2011. Several cases of boredom in the Preliminary study had already been observed, with one student actually falling asleep during the experiment. These cases all occurred in students in the control condition. Because note taking had been removed as an option in the Preliminary study, these students had no opportunity to take notes and received no instant message interruptions. Their only option was to sit passively watching a lecture, a task beyond the reach of a few students. Consideration of this phenomenon led to the role of cognitive arousal and differential learning effects associated with it.

Quiz scores from the Preliminary study were consistent with boredom or activation mediating the effect of instant message interruption. Two changes were made in the software to investigate this possibility. First, the LES was removed and the modified EPQ–BV scale put in its place. In addition, the ability for the student to take notes was added, so that presence or absence of note taking became an experimental variable. Consideration of the Yerkes–Dodson curve also led to the inclusion of an intermediate level of 12 interruptions in addition to the 6 and 18 used in the Preliminary study.

Design

The Concluding study used a 2×4×2 factorial crossed design, with instant message contents (related versus unrelated), interruption frequency (0, 6, 12 or 18 messages) and note taking (available or unavailable) as the between–subject factors. Note that because the control group received no messages, the instant message contents dimension is collapsed for this group. The different combinations of treatments (message relatedness, frequency and note taking) addressed experimental question 5 regarding interactions between factors.

Participants

Of the students assigned from the subject pool to participate in this study, 203 elected to participate in exchange for receiving experimental credit. Of these, six cases were excluded from analysis. In three cases, the data failed to record properly during the experiment, and no usable data were obtained. A fourth student reported after the experiment that she had previously taken the class the lecture was from. A fifth student was excluded because she responded to an instant message on her cell phone for several minutes during the experiment. The sixth student took the flash drive the information was recorded on after the experiment, making it impossible to include their information in the analysis. Analysis was conducted on the remaining 197 students, providing a cell size of 14 for most conditions.

The sample consisted of 137 female and 60 male students. Age ranged from 17 to 43, with an average of 20.85 and a standard deviation of 2.58. The median and modal ages were both 21. There were 15 freshman, 30 sophomores, 52 juniors, 99 seniors and one graduate student. Students came from an assortment of majors, with 25 Business, 28 Communication, 21 Education, 2 Engineering, 41 Liberal Arts, 55 Natural Sciences, 7 Nursing, 2 Social Work, and 16 Undeclared.

Procedure

As with prior studies, students watched the lecture, responding to instant messages when they appeared and taking notes on the lecture content, should they have the option and wish to do so (see Figure 2). Immediately at the end of the lecture. students were taken to the first page of the lecture quiz (see Figure 4), previously described in the Measures section.

Students worked through the three page quiz, and were then taken to the EPQ–BV scale (Figure 6). After finishing this, they concluded by answering demographic items (Figure 7). When finished, they were informed that their participation was complete and they were free to go (Figure 3).

4. Results

The results for each study are presented, after which results are compared across studies. The chapter concludes with an evaluation of the experimental hypotheses. All analyses were conduced using IBM's PASWStatistics 18.0.3 for Macintosh OS X. Dr. Matthew Hersh of the College of Natural Sciences supervised an independent analysis of the Pilot and Preliminary studies conducted by graduate students in the Division of Statistics and Scientific Computation. These researchers used SAS 9.2 to perform analyses using ANOVA, with results matching mine when outliers are included. Unless otherwise specified, all graphs representing percentages range from 0 to 100. Graphs representing scores on the recall quiz range from 0 to 10.

1. Pilot Study

Quiz

Initial review of the Pilot study data for outliers showed three points more than two standard deviations outside the mean score, one of which (subject 12) also stood out from scores within the group (see Figure 8). Analysis indicated that this score (z = -2.79) played a deciding role in determining significance within the study. Following the advice of Miles and Shevlin (2001) and Pedhazur and Schmelkin (1991), results are presented both with and without this outlier to avoid misleading the reader.

Outlier Included

A combined ANOVA across all five experimental groups failed to reach significance ($F(4,151) = 2.01$, $p = .096$, $MSE = 1.93$). Students responding to 18 or 36 lecture related or unrelated interruptions in a 24–min. lecture recalled the lecture items as well as students with no interruptions.

Because a crossed design was used, it was possible to aggregate data both with regard to number of interruptions and for related and unrelated messages (control excluded from analysis). An ANOVA comparison across number of interruptions did not reach significance ($F(2,153) = 2.51$, $p = .085$, $MSE = 1.93$), indicating no difference in recall regardless of interruption condition (0, 18 or 36).

A second ANOVA comparison across relatedness of interruptions also failed to reach significance ($F(1,123) = 3.27$, $p = .073$, $MSE = 1.80$), indicating no difference in recall if the messages were related or unrelated to the lecture. Note that the control condition was excluded from this analysis as it had no relatedness value.

These factors were combined to test for interaction, and did not reach significance ($F(3,121) = 1.12$, $p = .343$, $MSE = 1.81$). Note that the control condition was excluded from these analyses as it had no value for relatedness.

Outlier Removed

A combined ANOVA across all five experimental groups found a significant difference between groups ($F(4,150) = 2.87$, $p = .025$, $MSE = 1.86$). Post hoc analysis with Tukey's HSD indicated that this difference was due to poor performance by students receiving messages related to the lecture content. As shown in Figure 9, students who received either 18 or 36 lecture related messages did significantly worse than the control group on the ten–point lecture quiz, respectively scoring an average of 1.31 ($p < .05$) and 1.44 ($p < .024$) points lower than the control group. In other words, lecture–related interruptions impeded performance relative to the control group, but unrelated ones did not. The unrelated 18 and 36 interruption groups also scored lower, but the difference in the means from the control group was smaller and did not reach significance 0.86 ($p < .369$) and 0.73 ($p < .538$) respectively.

An ANOVA comparison across number of interruptions did reach significance ($F(2,152) = 4.09$, $p = .019$, $MSE = 1.86$), indicating a difference in recall due to interruption condition (0, 18 or 36). Additional analysis via Tukey's HSD (adjusted $p < .026$) indicated that the difference was due to poor performance by students in the 18 and 36 interruption conditions as compared with the control condition. Because the outlier was in the control group, a second ANOVA would have found no significant effect for relatedness of interruption.

These factors were combined to test for interaction, and did not reach significance ($F(3,121) = 1.12$, $p = .343$, $MSE = 1.37$). Note that the control condition was excluded from these analyses as it had no value for relatedness.

Interest and Multitasking self–ratings

In addition to the main experimental question, students were also surveyed about their own multitasking abilities and interest in the classroom lecture. Because these items were likely to be unrelated to score, data from the outlying case (subject 12) were included in these measures.

Interest in the clip was assessed on a seven–item Likert scale question, which asked "The video clip was chosen partly because it was expected to be of interest to education majors. How interesting was the content of the video clip to you?" Higher scores on this item reflect increasing interest. Thus a score of 1 would indicate no interest in the lecture, while a score of 7 would indicate extreme interest. The mean score was 4.54, slightly above the midpoint of 4 and negatively skewed (-.339). Students reported moderate to high levels of interest, with 9% selecting a 1 or 2, indicating little to no interest in the topic. An additional 65% selected 3 or 4, indicating moderate interest. The remaining 23% of students selected a 5 or higher, indicating a high level of interest in the topic. The professed interest in the lecture topic did not translate into better test scores, with a linear regression indicating no predictive value either including ($F(1,154) = 2.41$, $p < .122$, $MSE = 1.94$, ß = .124, adjusted $R^2 = .009$) or excluding the outlier ($F(1,153) = 1.75$, $p < .118$, $MSE = 1.90$, ß = .106, adjusted $R^2 = .005$).

To gain insight into beliefs about multitasking, a five–item Likert scale question asked “This experiment explored how well students can multitask – that is, attend to more than one thing at a time. What kind of impact do you think multitasking has on the number of mistakes you make?” Higher scores on this item reflect increasing errors. Thus a score of 1 would indicate no errors due to multitasking, whereas a score of 5 would indicate twice the number of errors. The mean score was 2.79, slightly below the midpoint of 3 and positively skewed (.627). In estimating their ability, 41% of respondents selected 1 or 2, indicating their multitasking would lead to little if any increase in error rates, 42% selected a moderate effect by selecting 3, and only 17% estimated a large increase in errors by selecting 4 or 5.

These results lend empirical support to informal observations that students are confident in their multitasking abilities (Bowman et al., 2010; Watson & Strayer, 2010).

However this confidence seems misplaced. A linear regression indicated that self–rated multitasking ability was a significant negative predictor of performance on the retention quiz, both with subject 12 included ($F(1,153) = 8.03$, $p < .005$, $MSE = 1.93$, ß = -.223) and without ($F(1,152) = 8.10$, $p < .005$, $MSE = 1.86$, ß = -.225). While it is uncertain why this association exists, the work of Kruger and Dunning (1999) may be applicable. These results are depicted in Figure 10.

2. Preliminary Study

Quiz

Initial review of the Preliminary study data for outliers showed one point (subject 90) stood out from scores within the group (see Figure 11). Analysis indicated that this score ($z = -3.05$) played a deciding role in determining significance within the study. Results are presented both with and without this outlier to avoid misleading the reader.

Outlier Included

A combined ANOVA across all five experimental groups (including the control group) failed to reach significance ($F(4,169) = 2.01$, $p = .095$, $MSE = 1.92$). Students responding to 0, 18 or 36 lecture related or unrelated interruptions in a 24–minute lecture recalled the lecture items as well as students with no interruptions (see Figure 12).

Because a fully crossed design was used, it was possible to aggregate data both with regard to number of interruptions and for related and unrelated messages (with the control group excluded from the analysis). Note taking was not an option in this experiment, so no analysis is possible. An ANOVA comparison across number of interruptions did not reach significance ($F(2,171) = 1.94$, $p = .148$, $MSE = 1.93$), indicating no difference in recall regardless of interruption condition (0, 6 or 18).

A second ANOVA comparison across relatedness of interruptions also failed to reach significance ($F(1,114) = 3.64$, $p = .059$, $MSE = 1.87$), indicating no difference in recall if the messages were related or unrelated to the lecture. These factors were combined to test for interaction, and did not reach significance ($F(3,112) = 2.48$, $p = .062$, $MSE = 1.86$). Note that the control condition was excluded from these analyses as it had no value for relatedness.

Outlier Removed

A combined ANOVA across all five experimental groups (including the control group) still reflected a significant difference between groups ($F(4,168) = 2.64$, $p < .035$, $MSE = 1.89$). Post hoc analysis with Tukey's HSD indicated that this difference was due to higher performance by students in one experimental group as compared with the control and other experimental groups. Specifically, students who received 6 unrelated messages did significantly *better* than both the control group and the experimental groups receiving 18 related messages. The mean score for this group (8.0) was over a point higher than the mean of the next highest group (6.9). The size of this effect was small (adjusted $R^2 = .037$). No other groups differed significantly.

An ANOVA comparison across number of interruptions did not reach significance ($F(2,170) = 2.48$, $p = .087$, $MSE = 1.90$), indicating a no significant difference in recall due to interruption condition (0, 18 or 36).

In contrast to the results including the outlier, a second ANOVA comparison across relatedness of interruptions did reach significance when the outlier was removed ($F(1,113) = 4.62$, $p < .034$, $MSE = 1.83$), indicating that students showed better retention with unrelated as opposed to related messages, though the size of this effect was small (adjusted $R^2 = .031$). These factors were combined to test for interaction. The results also contrasted with the outlier analysis, and reached significance ($F(3,111) = 3.34$, $p < .022$, $MSE = 1.81$). Post hoc analysis indicated that the interaction effect was due to better retention by students with 18 (but not 36) messages unrelated to the lecture. As before, the size of this effect was small (adjusted $R^2 = .058$). Note that the control condition was excluded from these analyses as it had no value for relatedness.

LES Scale

In discussion with R. Kay (personal communication, May 1, 2011), he noted “in our study, the prof was top notch and made considerable strides to engage the students. All students had laptops and connectivity” and that the bottom line appears to be "engage the students or they will use the laptop for anything but academic purposes.” The results of the survey indicate that students do vary their computer use in response to classroom demands, and are capable of self–regulation in the classroom.

Analysis of the modified LES scale (Lauricella & Kay, 2010) provided insight into student variations in computer use. As mentioned in the previous chapter, the scale uses a self–report of computer use in general and within the context of the specific class. Two academic uses (note taking and “other”) and four non–academic uses (email, instant messaging, playing games and web surfing) are measured. An academic and a non–

academic score were calculated by taking the mean of each student's response to those four questions. A linear regression analysis with these scores indicated a significant positive relationship between academic computer use in the classroom and performance on the recall quiz ($F(1,172) = 26.10$, $p < .008$, $MSE = 1.94$, ß = .200, adjusted $R^2 = .035$). In contrast non–academic use of computers in the classroom did not predict performance on the recall quiz ($F(1,172) = 3.253$, $p < .35$, $MSE = 1.91$, ß = -.071, adjusted $R^2 = -.001$). The distribution of academic and non–academic use is shown in Figure 13.

The survey also gives better insight into students' reported use of computers in the classroom. Note Taking is the most popular use by far, with nearly half (48%) of students reporting that they used computers for course–related note taking between 75 and 100% of the time for all classes. The next most popular category was 0% or "never," with students reporting that for all classes they never used computers for email (71%), instant messaging (59%), games (81%) or non–academic web surfing (43%). Following the 0% use, the next highest category for non–academic use was 1–25%. A small number of students reported email (17%) and games (16%) use at this frequency, with a higher proportion of these students at this level of instant messaging (29%) and non–academic web surfing (36%). These results are depicted in Figure 14, with Table 1 providing a summary of the percentages of students and how they spent their classroom time.

Instant Message Average and Maximum

Four factors relating to instant messaging were investigated. Students were questioned about the number of IMs that they sent and received in a typical class, as well as the maximum number that they had ever sent or received in any class. Visual inspection (see Figure 15) suggested that sent and received messages were highly correlated, with a regression analysis finding significant correlations for both the typical ($F(1,172) = 852.63$, $p < .000$, $MSE = 0.48$, ß = .912, adjusted $R^2 = .831$) and maximum ($F(1,172) = 270.73$, $p < .000$, $MSE = 0.68$, ß = .879, adjusted $R^2 = .772$) questions. This indicates that students will actively respond to messages they receive, both on "typical" and "heavy" messaging days.

Because the numbers are so close, the mean of sent and received instant messages will be used to describe use. In a typical class, 25% of students exchanged three or fewer instant messages while 45% exchanged between four and seven messages. The categories for eight to eleven, twelve to fifteen, and sixteen or more messages were 13%, 9% and 8% respectively. For maximum messages, 14% of students exchanged three or fewer instant messages while 27% exchanged between four and seven messages. The number drops again for eight to eleven (15%) and twelve to fifteen messages (19%), then spikes again at 16 or more (25%). Maximum numbers were bimodal, with 61% reporting 8 or more IMs received and 57% sending 8 or more. These results are depicted in Figure 15 and summarized in Table 2.

3. Concluding Study

Initial review of the Concluding study data for outliers showed six points (subjects 1, 8, 10, 78, 87 and 163), which stood out from scores within the group (see Figures 16 and 17). An ANOVA of experimental groups indicated that these scores played a deciding role in determining significance. Including the outliers changed the significance of the group analysis from ($F(13,183) = 1.42$, $p = .155$, $MSE = 1.72$) to

($F(13,177) = 1.93$, $p < .030$, $MSE = 1.62$). However, the outliers were not as extreme as in the Pilot and Preliminary study – around 2 standard deviations from the mean for their group (-1.84, -1.26, -1.84, +2.18, -2.34 and -2.45, respectively). The first three of these scores were from one experimental group, suggesting that they were accurate measures of student performance and should be retained. An ANOVA of experimental groups including the scores for subjects 1, 8 and 10 while eliminating those for 78, 87 and 163 was not significant ($F(13,180) = 1.67$, $MSE = 1.66$, $p = .070$). As a result, all outliers were retained for this analysis.

Quiz

A combined ANOVA across all 14 experimental groups (including the control groups) found no significant differences ($F(13,183) = 1.42$, $p < .155$, $MSE = 1.72$) The means are displayed in Figure 18.

Because a crossed design was used, it was possible to aggregate data with regard to note taking, number of interruptions and relatedness of messages (control excluded from analysis). An ANOVA comparison across number of interruptions failed to reach significance ($F(3,193) = .17$, $p = .914$, $MSE = 1.75$), indicating no difference in recall regardless of number of interruptions during the lecture (0, 6, 12 or 18).

An ANOVA comparison of interruption relatedness just reached significance ($F(1,167) = 3.94$, $p < .049$, $MSE = 1.71$), indicating that students showed better retention with related as opposed to unrelated messages. The size of this effect was very small (adjusted $R^2 = .017$). Note that the control condition was excluded from this analysis as it had no relatedness value. Given the exclusion of the three control group outliers and the marginal significance found in this analysis, this result will be treated as non–significant.

An ANOVA comparison across note taking groups reached levels of non–significance rarely seen in the academic literature ($F(1,195) = .01$, $p = .939$, $MSE = 1.74$), indicating no difference in recall regardless of student note taking.

A crossed design analysis compared all three experimental conditions (note taking, number of interruptions and relatedness), and did not produce significant results ($F(2,157) = 1.74$, $p = .179$, $MSE = 1.68$).

EPQ–BV

As is typical with college age students (Eysenck & Eysenck, 1994), the modified EPQ–BV measure had an extremely negative skew (skewness = -.789, std. error of skewness = .173), with half of the students scoring in the top 17% of the scale. This scale was divided into quartiles and checked for significance against the scores of students based on frequency of interruption and the presence or absence of notes.

An ANOVA showed that cognitive arousal group was not related to quiz score ($F(3,197) = .24$, $p = .867$, $MSE = 1.75$), nor did it reach significance when combined with number of interruptions ($F(9,181) = .614$, $p = .784$, $MSE = 1.78$), note taking ($F(3,189) = 1.03$, $p = .379$, $MSE = 1.75$), or both ($F(9,165) = 2.09$, $p = .756$, $MSE = 1.80$).

Instant Message Average and Maximum

Four factors relating to instant messaging were investigated. Students were questioned about the number of IMs that they sent and received in a typical class, as well as the maximum number that they had ever sent or received in any class. Visual inspection (see Figure 19) suggested that sent and received messages were highly correlated, with a regression analysis finding significant correlations for both the typical (F(1,195) = 223.44, p < .000, MSE = 0.60, ß = .871, adjusted R^2 = .757) and maximum (F(1,195) = 723.18, p < .000, MSE = 0.64, ß = .887, adjusted R^2 = .787) questions. This indicates that students will actively respond to messages they receive, both on "typical" and "heavy" messaging days.

Because the numbers are so close, the mean of sent and received instant messages will be used to describe use. In a typical class, 16% of students exchanged three or fewer instant messages while 46% exchanged between four and seven messages. The categories for eight to eleven, twelve to fifteen, and sixteen or more messages were 17%, 9% and 13% respectively. For maximum messages, 9% of students exchanged three or fewer instant messages while 27% exchanged between four and seven messages. The number drops again for eight to eleven (19%) and twelve to fifteen messages (13%), then spikes again at 16 or more (32%). Maximum numbers were bimodal, with 36% reporting seven or fewer IMs per class and 64% with 8 or more. These results are depicted in Figure 19 and summarized in Table 2.

4. Combined Analysis of Preliminary and Concluding Studies

Initial review of the Concluding study data for outliers showed three points, which stood out from scores within the group. An ANOVA of experimental groups indicated that these scores did not affect significance. Therefore, all outliers were retained for this analysis.

The software interface in the Pilot study differed substantially from the version used in the Preliminary and Concluding studies, making inclusion of Pilot study data questionable. This analysis only included data from the Preliminary and Concluding studies, and notes where differences in cell sizes were compounded by the return of note taking as a variable and a twelve–interruption condition in the Concluding study. The numbers of cases per condition are shown in Table 3.

Quiz

For number of interruptions, the control groups provided 86 cases, the six IM condition had 115, the twelve IM condition had 56 and the eighteen IM condition had 114. Results for number of interruptions were non–significant ($F(3,367) = .88$, $p = .454$, $MSE = 1.84$). Removal of the small 12 IM condition from the analysis had no appreciable effect on significance.

With the control group excluded, the relatedness condition cells were very close – 147 and 138 cases respectively. Consistent with the previous results, the means for the two groups were also very close, and did not differ significantly ($F(1,283) = .04$, $p = .845$, $MSE = 1.80$).

Comparing sample sizes, the note taking groups had 273 cases while the no–notes groups had 98. As before, a comparison of the two conditions was non–significant ($F(1,369) = .15$, $p = .702$, $MSE = 1.84$), though it should be noted that this result may be biased from the nearly 3:1 ratio of sample sizes.

A final ANOVA checked for interactions between all experimental groups (note taking, number of interruptions and relatedness), and produced non–significant results ($F(13,357) = 1.57$, $p = .092$, $MSE = 1.82$).

Instant Message Average and Maximum

Four factors relating to instant messaging were investigated. Students were questioned about the number of IMs they sent and received in a typical class, as well as the maximum number they had ever sent or received in any class. Visual inspection (see Figure 19) suggested that sent and received messages were highly correlated, with a regression analysis finding significant correlations for both the typical ($F(1,369) = 1413.94$, $p < .000$, $MSE = 0.55$, ß = .891, adjusted $R^2 = .792$) and maximum ($F(1,369) = 1295.57$, $p < .000$, $MSE = 0.66$, ß = .882, adjusted $R^2 = .778$) questions. This indicates that students will actively respond to messages they receive, both on "typical" and "heavy" messaging days.

Because the numbers are so close, the mean of sent and received instant messages will be used to describe use. In a typical class, 21% of students exchanged three or fewer instant messages while 45% exchanged between four and seven messages. The categories for eight to eleven, twelve to fifteen, and sixteen or more messages were 15%, 9% and 11% respectively. For maximum messages, 12% of students exchanged three or fewer instant messages while 27% exchanged between four and seven messages. The number drops again for eight to eleven (17%) and twelve to fifteen messages (16%), then spikes again at 16 or more (29%). Maximum numbers were bimodal, with 66% reporting seven or fewer IMs per class and 34% with eight or more.

5. Discussion

Research Question 1

It was hypothesized that larger numbers of messages would produce greater processing deficits, resulting in lower scores on the recall quiz. This hypothesis took the form of the question, “Does the frequency of IM interruption have an impact on lecture recall?”

Cognitive models of information processing such as Baddeley’s suggest that there should be a significant negative linear relationship between number of instant messages and lecture retention due to limitations on cognitive processing capacity. However, the Yerkes–Dodson model predicts that an inverted “U” curve would emerge in response to an increased number of instant message interruptions. The literature on student multitasking presents a more complex picture, with both significant and non–significant results reported. Results from the study were similarly mixed. The Pilot study reached significance only with outliers removed, all other results for interruption frequency were non–significant.

There are several possible explanations for this result. For example, it is possible that the redundancy of lecture presentation, with essential information repeated several times throughout the lecture. In preparing the interruption timing, it was noticed that all topics covered in the lecture retention quiz were mentioned at least twice in the video lecture used in this experiment, and most were repeated three or more times. Unfortunately testing for the effect of lecture redundancy is beyond the scope of this research, and will have to be tested in a separate experiment. Whatever the reason, the result is the same – in this experiment student scores on the lecture retention quiz were not affected by the number of instant messages they received. Accordingly the hypothesis is rejected.

Research Question 2

It was hypothesized that instant messages unrelated to the lecture would produce greater processing interference than related messages. This was expected to result in lower scores on the recall quiz for unrelated messages than related ones. This hypothesis took the form of the question, "Do IM questions relatedness to the lecture have an impact on lecture recall?"

The relatedness results from the Pilot study were not significant, while the relatedness results from the Preliminary and Concluding studies were statistically significant. However these relationships were in opposite directions for the two studies. Students in the Preliminary study did better with interruptions unrelated to the lecture, while those in the Concluding study did better with interruptions related to the lecture. When the Preliminary and Concluding studies were combined for analysis the effects cancelled each other out, leaving a non–significant result.

Similarly, the marginally significant results from the Concluding study were consistent with the hypothesis that related messages improve scores. Observations of student boredom during the Preliminary study led to the hypothesis that scores improved due to the interruptions providing a break from the lecture topic, increasing student arousal and relieving boredom. However, these results are inconsistent with those from the Preliminary and Pilot studies. Thus relatedness results are unclear from this experiment.

Research Question 3

It was hypothesized that taking notes will result in higher scores on the recall quiz. This hypothesis took the form of the question, "Does the opportunity to take notes affect lecture recall?"

The research literature supports the idea that note taking contributes to student success, and this study examined whether the activity of taking notes allows students to fight boredom by providing a source of cognitive arousal. This result was not observed in this experiment. Direct comparison of the notes versus no–notes conditions in the Concluding study was not significant, nor were significant results observed in the combined analysis of Preliminary and Concluding studies. The difference between these results and those of other researchers may be due to the differences between a single experimental session and the extended note taking typical of the regular semester. In other words, this design may not be an appropriate measure of how taking notes impacts lecture retention in the classroom.

Taken together as a whole, these results do not support the hypothesis that taking notes with a computer promotes lecture retention by allowing bored students to increase their arousal level.

Research Question 4

It was hypothesized that students with lower cognitive arousal would show recall quiz benefits from interruptions, while those with high arousal would have impaired performance. This hypothesis took the form of the question, “Do cognitive arousal traits affect lecture recall?”

Underlying differences in cognitive arousal were expected to result in differential responses to interruptions, with high arousal students (introverts) scoring lower on lecture recall as number of interruptions increases, regardless of relevance. Low arousal students (introverts) would be expected to show improvement with interruption until they reach their optimum level of arousal, after which performance should decrease. This was not observed, with no difference in performance found between the two groups in the Concluding study. Unfortunately, the cognitive arousal scale was a late addition to the experiment, and was not administered in the Pilot or Preliminary study.

There are several possible explanations for this result. The scale used in this experiment (the EPQ–BV) has not been used as extensively as the EPQR–S, and is traditionally administered as a pencil–and–paper instrument rather than in an online format. As a result, the test used in this experiment may not have provided an accurate rating of cognitive arousal. It is also possible that the interruptions were not arousing enough to differentially affect the performance of the groups. Given the highly skewed results from this instrument, a Type II error due to restriction of range is also a possibility. Approximately half of the students surveyed scored 85% or higher on this measure. Whatever the reason, this experiment did not support the hypothesis that cognitive arousal is a significant factor in lecture retention.

Research Question 5

It was hypothesized that interaction effects beyond those of individual factors would appear as measured by scores on the recall quiz. This hypothesis took the form of the question, “Are there any significant interactions between these factors?”

No significant interactions involving note taking were found. Interaction effects were observed in the Preliminary study when outliers were removed. In this case students with six unrelated messages scored significantly better than those with six related messages. A similar comparison of the 18–message condition found no significant difference between related and unrelated. A significant result was not found in either the Pilot or Concluding studies. The hypothesis is not supported in this experiment, there appears to be no significant interaction between number and relatedness of interruptions.

Research Question 6

It was hypothesized that students in this experiment would report computer use consistent with this pattern, leading to the question, “How do students use computers in classroom situations?”

The LES scale, along with student instant messaging self–report, addressed this question. The results suggest that academic use of laptops was a small but significant predictor of quiz score, while non–academic use did not reach significance.

As described in the results section, the most common use of computers in the classroom was for note taking, with 48% of students reporting this use between 75 and 100% of class time, and substantial numbers stating that they never use computers for non–academic activities in the classroom. Although there are students who use computers for non–academic activities in the classroom, these appear to be the minority. When considering bans on laptops in the classroom, the instructor may wish to consider these results when deciding what appropriate levels of computer use are. From the data, it appears that academic use is at the core of classroom laptop activity.

Research Question 7

It was hypothesized that students who estimated they have higher multitasking proficiency would have lower scores on the recall quiz than those who do not claim such proficiency. This hypothesis took the form of the question, “How accurate is student self–assessment of multitasking ability?”

Like the student who claimed to be able to engage in five overlapping tasks (Lenhart et al., 2001), some people have an inflated view of their ability to multitask. However, quantifiable measures of this belief have not been used to show their magnitude and how common they are. A significant relationship between self–ratings of multitasking and recall quiz score emerged when students in the Pilot study were asked to estimate the impact that multitasking would have on making mistakes while doing homework. Only a minority of 17% felt that trying to multitask would lead to a large increase in errors. The remaining students felt there would be little or moderate impact. This provides empirical support for the assertion that students feel comfortable multitasking and anticipate little down–side in doing so. As mentioned in the results section, student self–rating for multitasking performance did not match with recall quiz scores. In fact, there was a strong negative correlation between self–rating and final score. It appears that student estimates of multitasking ability were good negative predictors of exam score.

Limitations

Despite the effort to provide an experimental environment similar to the classroom, there were several limitations of the current study. At 24 minutes, the lecture was about half the duration of a typical classroom lecture. In addition there was only a single lecture rather than a series of lectures as would occur in a typical class. It could also be argued that the subject matter was not inherently interesting, that students in the subject pool have low motivation to perform well, or that the instant messaging interruptions needed to be more engaging and conversational. While reported level of interest in the subject was moderate to high, a few students exhibited obvious signs of boredom. Again the underlying reason for the boredom is unknown, suggesting an actual class could provide a more ecologically valid setting for research, provided the technological, logistical and ethical issues involved could be addressed. While most students today do have laptops, ensuring all students in a class have them and have functioning instant message accounts and software is not only difficult, but has the potential to cut into limited classroom time. Similarly, though the results of this experiment indicate this would not be the case, random assignment of students to conditions may raise ethical issues due to a potential impact on classroom grades. The issue of timing the interruptions so lecture–relevant messages occur at the proper time would also be complex.

An additional concern is the type of messaging used in this experiment. Shiu and Lenhart (2004) found that 29% of 18 to 27 year olds engaged in multiple simultaneous IM conversations on a daily basis. This experiment did not address the possibility of requiring students to keep track of single or multiple conversations, merely to briefly respond to one type of interrupting message. Given the frequency with which student engage in multiple conversations, and the expected increase in attentional requirements, a greater impact on lecture recall would be anticipated if this component were added to the design. Unfortunately this would require either a team of experimenters to actively interact with students, or the use of advanced natural language processing such as Apple's Siri software for the iPhone. Either option would require a considerable investment in terms of personnel and software.

The recall quiz was also of some concern. It was taken from an actual classroom quiz, so it is ecologically valid. However, on a ten–point scale across all three studies, only a single student recorded score of one, and none scored a zero. In contrast there were 84 individuals who scored nine points on the recall quiz and 38 who scored ten. This negatively skewed scoring raises statistical concern. A refined measure with a more normal curve would provide a less biased estimate of student lecture retention due to restriction of range in the scores. It should also be noted that the recall quiz occurred immediately after the lecture, and that delayed recall was not tested in this research.

Finally, the level of analysis in the study was at the level of groups of students. It would have been possible to increase group size by testing multiple conditions within individual students, so half the time was spent with related messages, and the other half was spent with unrelated messages. This should reduce the effect of variance between students, increasing the power of the experiment. Under these circumstances the level of cognitive arousal would remain constant within the subject, reducing the problem of skewed distribution found with the scale used in the Concluding experiment.

General Implications

In doing research for this study and in interpreting the results, it has become clear that studies of multi–tasking in students have yielded mixed findings, with non–significant results appearing in much of the literature. These experiments fell into two categories, those requiring multitasking either within a single Baddeley processing store (visuospatial or phonetic), and those requiring processing in both stores. As shown in Table 4, multitasking had an effect on task performance speed if either store is used. However, the same experiments found significant differences on student scores only where multitasking required simultaneous phonological and visuospatial processing. The non–significant results of this study are consistent with this pattern, as it required multitasking only within the phonological memory store.

Within the classroom, the most common academic use of computers is note taking. Combined with lectures, this requires multitasking within the phonological store, but not the visuospatial one. Similarly the most popular (though less widely adopted) non–academic uses are instant messaging and web surfing, both of which make low demands on the visuospatial store due to the automaticity of reading. Regardless of the reason, the lack of significant performance deficits in this and previous studies indicates that the rapid embrace of laptop computers has not significantly affected college student success.

This study also adds to the literature on how students are currently using computers in the classroom. Self–reports indicate most students spend the majority of time in course–related computer activities, with non–related activities occurring far less frequently. Instructors who are concerned their students are not using their time effectively should be reassured by these results.

Finally, it is not surprising that those students who rated themselves as better at multitasking did significantly worse on the recall quiz than those who rated themselves lower. This inflated estimate of ability may reflect students lacking the expertise to recognize their own lack of ability, as described by Kruger and Dunning (1999). While research is needed to evaluate this possibility, instructors should be aware that student estimates of multitasking may be inflated, and resistant to correction.

In summary, this research suggests that at present both concerns and promises about the impact of computers in education are overblown, and students will continue to succeed in college, regardless of the presence or absence of technology.

Educational Implications

When faced with the issue of managing laptop use in the classroom, it appears that instructors have little to worry about if their sole concern is student grades. Note taking, message relevance and number of interruptions failed to have significant effects individually or in combination. Despite the lack of significant results, some instructors and students will still be concerned about wireless technology in the classroom. While the majority of students were unaffected by the interruptions, some students and instructors regularly report being distracted by other student's computer use. In addition, devoting attention to computers can impact student interaction and discussion.

Instructors can adjust by adopting strategies to promote effective computer use in the classroom. For example, McCreary (2009) instituted a partial ban, explaining "I ban laptops –from the first few rows of my classroom only. I do this because I recognize that some students use laptops effectively and appropriately and benefit from having them in the classroom. I do this because other students cannot resist the temptation to look at another student's screen and therefore need a place to sit in the classroom from distraction." In a similar vein, instructors may encourage students to participate in classroom activities. Murray (2011) notes that instructors, "could make class participation a portion of the final grade, offer grade boosts for those who participate in class most often and best, or offer transcript notations at law schools that provide such designations."

While an outright ban on laptops in the classroom will always appeal to some individuals, instructors should be reassured by the results of this study that the laptops were not harming student learning as measured in this experiment.

Research Directions

Considering the non–significant results for personality, and the similarities between this experiment and previous ones that included cognitive arousal as a factor (Furnham et al., 1994; Furnham & Bradley, 1997; Gonder, 2010), a new hypothesis should be tested. Specifically, that cognitive arousal only becomes a factor in recall when subjects multitask using both phonological and visuospatial systems. Unfortunately the need for this hypothesis only became apparent after the results of the Concluding study were compared with the experimental literature. Additional research using a mix of visual and auditory processing stores in association with cognitive arousal may serve to clarify the conflicting results of this and prior experiments.

While not a formal part of the study, unanticipated experimental observations play an important part in suggesting hypotheses and future experiments. In this case, two informal observations were notable. In the first case, a student in the preliminary experiment had been assigned to the control condition. Accordingly, they received no instant message interruptions and could not take notes. Their sole task was to watch the 24–minute lecture clip. This student leaned back in their chair and fell asleep, waking when the lecture stopped to respond to the exam. He felt this was appropriate, saying "The subject matter was irrelevant to me and I was completing the study on the last day of an extremely difficult and long week of tests, projects, and work" and "I was never asked any questions through the 'instant messaging' system that was in place." This student's data was excluded from analysis. These observations lead to the consideration of cognitive arousal as a mediating factor, which should be examined in the follow–up experiment. This suspicion was supported by a similar incident in the follow–up study. A student with her back to me was in the same experimental condition as the previous student, with no interruptions or note taking. While she listened to the lecture throughout, she spent the last five minutes of the lecture using the mouse to drag the experimental window (Figure 2) around the edges of the screen as rapidly as she could. Data from this student was included in the analysis.

Research into the cognitive arousal trait is less common today than during the 1960s and 1970s, but these observations suggest the construct should be considered when research or lecture includes a passive lecture condition. The arousal scale used in this experiment was highly skewed, with 47% of students rating themselves in the top 15% of the revised cognitive arousal (extraversion) sub–scale (Eysenck, Eysenck, & Barrett, 1985).

Inclusion of Eysenck's (1976) cognitive arousal (extraversion) factor was a late addition to the experimental design, and is the most problematic aspect of this research for two reasons. First, the version of the scale that was used was not the full scale, but a subset of the original. In future experiments a better–validated measure may be best. A good contender for this would be the short form version of the EPQ, the EPQ–SF. This instrument is short enough to fit well with the computer–based test, uses more colloquial phrasing for the questions (Richendoller & Weaver, 1994; Weaver, 1991), and the scale validity is comparable (.81) to that of the EPQ–R (Weaver & Kiewitz, 2007).

Finally, there is the potential to use self–rated multitasking ability as a diagnostic tool for identifying students that may need help in developing a more realistic understanding of their academic abilities and limitations.

Appendix A

Related and unrelated messages used as interruptions

Related Messages

Is the eye contact method appropriate for middle school and older students?
Is social discomfort the basis of the eye contact method?
Does using the eye contact method depend on having imposing physical size?
Does assertive body language demonstrate aggression or anger?
Does assertive body language demonstrate submission or deference?
Does social discomfort mean that you don't care that people are judging you negatively?
Is the disruptive student hoping to annoy their classmates?
Is reviewing and explaining the classroom rules part of the eye contact method?
Is seeing and being able to get close to students a part of the eye contact method?
Is giving students their own "turf" part of the eye contact method?
Is arranging and moving around in the classroom important for the eye contact method?
Is catching disruptions early essential to the eye contact method?
Is consistent use of the eye contact method necessary for success?
Is a good teacher sensitive to patterns of student behavior as part of classroom management?
Is showing an emotional reaction to student disruptions essential for the eye contact method?
Is turning to look at the disruptive student the first step in the eye contact method?
Should you fold your arms in front of you when you make eye contact with a student?
Does the eye contact method require you to know the names of your students?
Should you use an angry tone when you say the disruptive student's name?
Should you use a straight, flat tone when you say the disruptive student's name?
Do you assertively walk to the disruptive student's location if they don't respond to their name?
Is staying out of the "personal space" of a disruptive student part of the eye contact method?
Are saying "good" and maintaining eye contact for 3 seconds both parts of the eye contact method?
Is physical contact with the student a part of the eye contact method?
Is removing a disruptive student from the classroom part of the eye contact method?
If a student resumes their disruption, should you repeat all the steps of the eye contact method?
Is dealing quickly with disruptions (overlapping) a feature of the eye contact method?

Unrelated Messages

Did you watch American Idol last week?
Did you watch Dancing With The Stars last week?
Is this your last semester at UT?
Was UT your first choice for college?
Have you ever gone to see a UT football game outside of Texas?
Have you gone to see a live music performance this semester?
Do you regularly update you Facebook profile?
Have you ever been up on the top of The Tower?
Have you ever visited any of UT's museums?
Will you stay inside of Texas for winter break?
Will you go outside of Texas for winter break?
Have you ever gone to see a UT football game outside of Austin?
Have your allergies been really bad this semester?
Are you planning to move outside of Texas after you graduate?
Are you planning to stay in Texas after you graduate?
Did you live in Austin before you started school here?
Do you know anyone who is loosing their job because of budget cuts?
Did you go to South by Southwest Music Festival this year?
Did you go to Austin City Limits Fest this summer?
Will you go to Eeyore's birthday party in Pease Park this spring?
Will you go to see the Trail of Lights over winter break?
Have you ever attended a tailgate party?
Have you watched the bats at sunset from the Congress Street Bridge?
Have you changed majors since you started at UT?
Do you plan to go to Sixth Street for the Halloween costume parade?
Have you ever missed class because you couldn't find parking?
Do you know the story of how Bevo got his name?

Appendix B

Items used in the eye–contact method quiz.

1. Which of the following is not a necessary precondition for the eye–contact method? a. Catching disruption early (wittiness). b. Moving around the room. c. Explaining the rules. d. Arranging the room. e. Attending to two things at once (overlapping).
2. Which of the following is not an appropriate step in the eye–contact method? a. Placing palm on student's desk. b. Extending hand towards student. c. Saying the student's name. d. Tapping student on shoulder to make eye contact. e. Saying "Good" when disturbance ends.
3. "Man, that Mrs. Robinson must have eyes in the back of her head," Eddie comments to John. The classroom management concept most closely related to this description is a. Withitness b. Overlapping c. Momentum d. Smoothness
4. A student is chronically disruptive, and you finally decide that you have to remove the student from the classroom. The student refuses to leave and says that you can't make him go. Your best course of action is to a. Physically remove the student and send him to the principal. b. The student has the choice to leave or go – apply your class rules. c. Tell the student leave in a straight, flat tone until the student complies. d. Get help. Teachers are not responsible for dealing with defiant students.
5. Which of the following is true regarding the eye–contact method? a. It does not involve use of intimidation or physical force. b. It does not involve use of social discomfort or intimidation c. It does not involve use of verbal behavior, simply assertiveness. d. It does not involve use of physical force or proximity.

6. The eye–contact method is based on the assumption that most misbehavior in the classroom is motivated by a. A desire to be assertive. b. A desire to get out of schoolwork. c. A desire to look good in front of peers. d. A desire to cause the teach social discomfort.
7. Withitness, as a classroom management concept, refers to a. The teacher coming prepared and staying on task. b. The student coming prepared and staying on task. c. Having a well–controlled classroom environment. d. Attending to disruptions quickly and decisively.
8. After the teacher stops writing on the board and waits three–seconds, what is the next step if the disturbance does not end? a. Say the student's name in a straight, flat tone. b. Stand with arms at the side, then extend hand. c. Turn and face the student. d. Walk back to the student.
9. Which of the following is a reason why you should not place your hand on the student’s shoulder during a confrontation? a. The student may misinterpret the action as a sexual advance. b. The student may be angry and react physically. c. It is inappropriate to make physical contact with students. d. It is difficult to establish eye contact.
10. Overlapping, as a classroom management concept, refers to a. Dealing with multiple disruptions by different students at once. b. Dealing with multiple disruptions by the same student at once. c. Using nonverbal cues to handle misbehavior. d. Pausing to handle disruptions in class.

Appendix C

Items of the LES scale. Modified with permission from Lauricella and Kay, 2010.

Course Related Computer Use
1. How much of the lecture time in this course do you use the laptop to take notes or follow the lecture? a. 100–76% of the lecture time b. 75–51% of the lecture time c. 50–26% of the lecture time d. 25–1% of the lecture time e. 0% of the lecture time
2. How much of the lecture time in this course do you use the laptop for academic purposes relating to this class (i.e., following lecture, doing in–class assignments or activities, viewing course outline, etc.)? a. 100–76% of the lecture time b. 75–51% of the lecture time c. 50–26% of the lecture time d. 25–1% of the lecture time e. 0% of the lecture time
3. How much of the lecture time in other courses do you use the laptop to take notes or follow the lecture? a. 100–76% of the lecture time b. 75–51% of the lecture time c. 50–26% of the lecture time d. 25–1% of the lecture time e. 0% of the lecture time
4. How much of the lecture time in other courses do you use the laptop for academic purposes relating to those classes (i.e., following lecture, doing in–class assignments or activities, viewing course outline, etc.)? a. 100–76% of the lecture time b. 75–51% of the lecture time c. 50–26% of the lecture time d. 25–1% of the lecture time e. 0% of the lecture time

Non–Academic Computer Use

5. How much of the lecture time in this course do you use the laptop for email of any kind (Hotmail, Yahoo, Gmail, etc.) for purposes other than this course?
   a. 100–76% of the lecture time
   b. 75–51% of the lecture time
   c. 50–26% of the lecture time
   d. 25–1% of the lecture time
   e. 0% of the lecture time

6. How much of the lecture time in this course do you use the laptop for instant messaging?
   a. 100–76% of the lecture time
   b. 75–51% of the lecture time
   c. 50–26% of the lecture time
   d. 25–1% of the lecture time
   e. 0% of the lecture time

7. How much of the lecture time in other courses do you use the laptop for email of any kind (Hotmail, Yahoo, Gmail, etc.) for purposes other than this course?
   a. 100–76% of the lecture time
   b. 75–51% of the lecture time
   c. 50–26% of the lecture time
   d. 25–1% of the lecture time
   e. 0% of the lecture time

8. How much of the lecture time in other courses do you use the laptop for instant messaging (MSN, etc.)?
   a. 100–76% of the lecture time
   b. 75–51% of the lecture time
   c. 50–26% of the lecture time
   d. 25–1% of the lecture time
   e. 0% of the lecture time

9. How much of the lecture time in this course do you use the laptop to play games?
   a. 100–76% of the lecture time
   b. 75–51% of the lecture time
   c. 50–26% of the lecture time
   d. 25–1% of the lecture time
   e. 0% of the lecture time

10. How much of the lecture time in other courses do you use the laptop to play games?
a. 100–76% of the lecture time
b. 75–51% of the lecture time
c. 50–26% of the lecture time
d. 25–1% of the lecture time
e. 0% of the lecture time

11. How much of the lecture time in other courses do you use the laptop surf the web for purposes other than this course?
a. 100–76% of the lecture time
b. 75–51% of the lecture time
c. 50–26% of the lecture time
d. 25–1% of the lecture time
e. 0% of the lecture time

12. How much of the lecture time in other courses do you use the laptop for surf the web for purposes other than this course?
a. 100–76% of the lecture time
b. 75–51% of the lecture time
c. 50–26% of the lecture time
d. 25–1% of the lecture time
e. 0% of the lecture time

Appendix D

Supplemental questions included in the demographics portion of the Preliminary and Concluding studies. The first item in the list, item 5, was changed between the studies. The Preliminary study item asked about how similar the lecture was to standard classroom lectures. The Concluding study item asked how interesting the lecture was.

5. The video clip was chosen partly because the lecture format was expected to be a familiar format to undergraduates. How similar was the classroom lecture in the video to lectures you've attended in you classes?

Unlike any class I have taken.	Like about half the classes I have taken.	Like every class I have taken.

5. The video clip was chosen partly because the content is typical for a university lecture. But interest in lectures varies. How interested were you in the subject of this lecture?

Absolutely no interest at all	Meh - kind of interesting	The lecture was really interesting

6. Thinking over the last semester, if you used chat or instant messaging, how many texts or instant messages would you get during a typical class?
   a. 0 messages
   b. at least 4 messages
   c. at least 8 messages
   d. at least 12 messages
   e. 16 or more messages

7. Thinking over the last semester, if you used chat or instant messaging, how many texts or instant messages would you send during a typical class?
   a. 0 messages
   b. at least 4 messages
   c. at least 8 messages
   d. at least 12 messages
   e. 16 or more messages

8. Messaging use varies. Thinking over the last semester, what were the most texts or instant messages you got during a class?
   a. 0 messages
   b. at least 4 messages
   c. at least 8 messages
   d. at least 12 messages
   e. 16 or more messages

9. Messaging use varies. Thinking over the last semester, what were the most texts or instant messages you sent during a class?
   a. 0 messages
   b. at least 4 messages
   c. at least 8 messages
   d. at least 12 messages
   e. 16 or more messages

Table 1

Activities (in–class and general)		Percentage of class time				
		0%	**1–25%**	**26–50%**	**51–75%**	**76–100%**
Academic	% Take Notes	12, 12	8, 13	5, 10	8, 17	68, 48
	% Academic activities	16, 14	18, 23	16, 20	18, 18	33, 24
Non–academic	% Use email	66, 71	22, 17	7, 5	4, 5	2, 2
	% Instant messaging	39, 59	41, 29	13, 8	6, 3	1, 1
	% Play games	93, 81	6, 16	1, 2	1, 2	0, 0
	% Surf web	45, 43	35, 36	11, 12	8, 9	1, 1

Self–reported classroom computer activities for class with subject pool requirement (first number) and classes overall (second number). Rounded to nearest integer.

Table 2

Preliminary Study		Ave. In	Ave. Out	Max In	Max Out
N	Valid	174	174	174	174
	Missing	0	0	0	0
Mean		5.33	4.99	8.74	8.37
Mode		4	4	16	4
Std. Deviation		4.698	4.676	5.690	5.694
Variance		22.073	21.861	32.381	32.419
Skewness		.927	1.000	-.061	.019
Std. Error of Skewness		.184	.184	.184	.184
Kurtosis		.099	.210	-1.377	-1.399
Std. Error of Kurtosis		.366	.366	.366	.366
Range		4	4	4	4
Concluding Study		**Ave. In**	**Ave. Out**	**Max In**	**Max Out**
N	Valid	174	174	174	174
	Missing	0	0	0	0
Mean		5.33	4.99	8.74	8.37
Mode		4	4	16	4
Std. Deviation		4.698	4.676	5.690	5.694
Variance		22.073	21.861	32.381	32.419
Skewness		.927	1.000	-.061	.019
Std. Error of Skewness		.184	.184	.184	.184
Kurtosis		.099	.210	-1.377	-1.399
Std. Error of Kurtosis		.366	.366	.366	.366
Range		4	4	4	4

Summary of statistics from Preliminary and Concluding studies regarding instant messaging during class. Students could choose a response indicating 0, 4+, 8+, 12+ or 16+ messages during an average class. The columns are average number of messages received (Ave. In), average number of messages sent (Ave. Out), highest number of messages received (Max In), and highest number of messages sent (Max Out).

Table 3

Notes	Control	Relatedness	Interruption Frequency		
			6	12	18
Yes	14	Related	14	14	14
		Unrelated	14	14	14
No	72	Related	41	14	41
		Unrelated	46	14	45

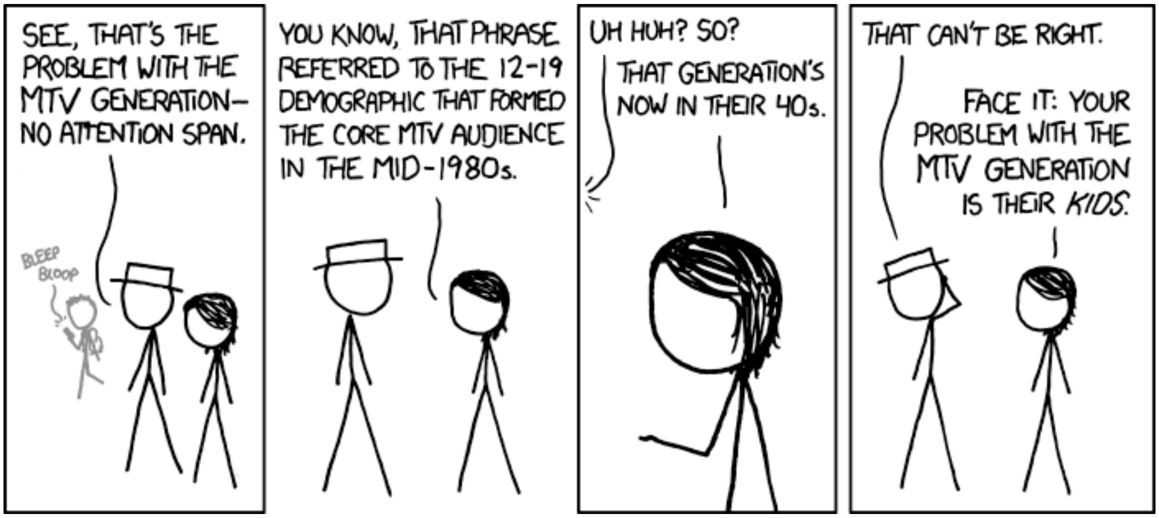
Number of cases per condition when Preliminary and Concluding studies are combined. Because notes and a 12 message condition were not used in the Preliminary study, these cell sizes are much larger than the others.

Table 4

Research	Modality		Impact	
	Unitary	**Mixed**	**Speed**	**Score**
Armstrong & Greenberg, 1990		X	X	3/7
Bowman et al., 2010	X		X	
Fox, Rosen, & Crawford, 2009	X		X	
Hembrooke & Gay, 2003		X		X
Pool, van der Voort, Beentjes, & Koolstra, 2000 (Experiment 2)		X	X	

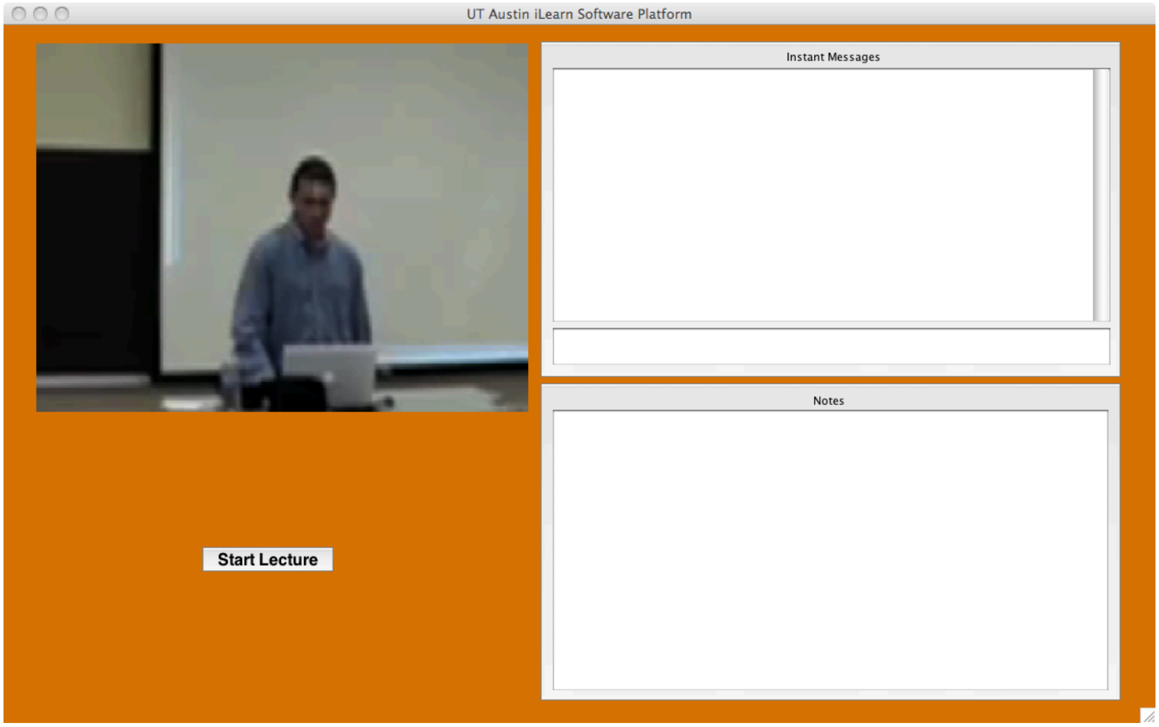
A comparison of multitasking studies by modality and effect. The Armstrong study found significant results in three tests (Nelson–Denny, Towers of Hanoi and creativity) while non–significant results were found for four tests (digit span, mental arithmetic, sentence verification and a letter series task). The Pool study utilized two experiments, the first of which did not require the students to switch attention to between the television set and the assignment. Since this does not meet the definition of multitasking used here, only results for the second experiment are reported.

Figure 1



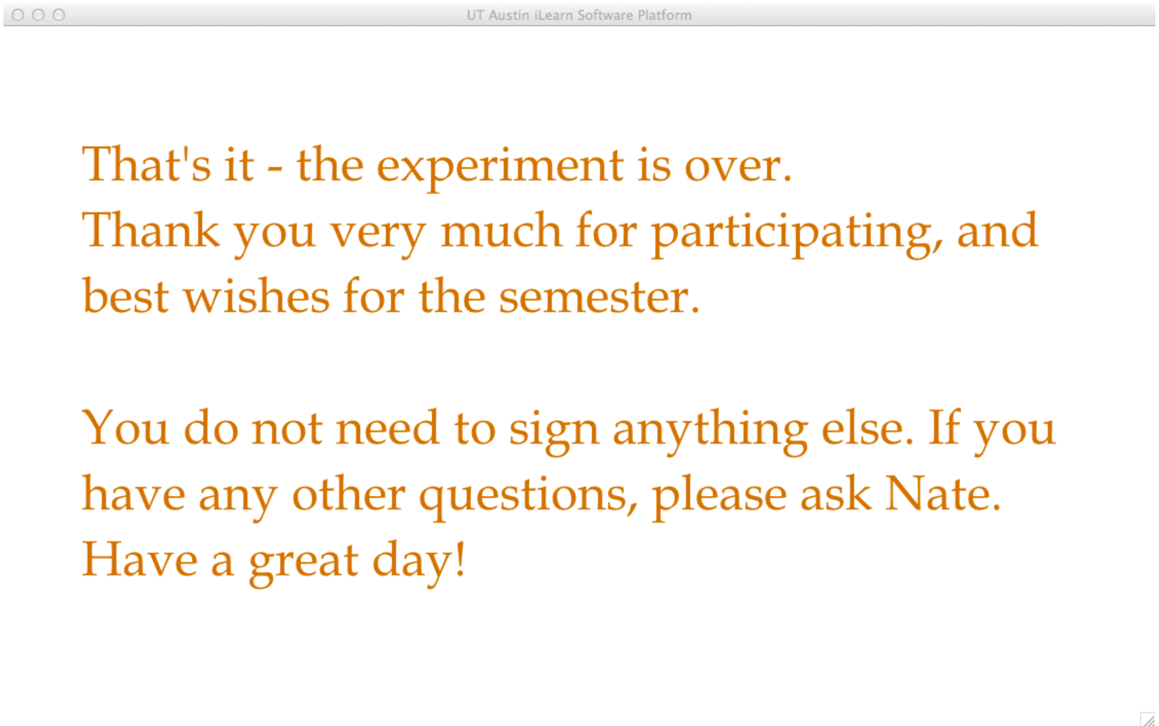
One example of the impact of technology on today's students and instructors, more commonly known as "Get off my lawn syndrome." (Munroe, 2011).

Figure 2



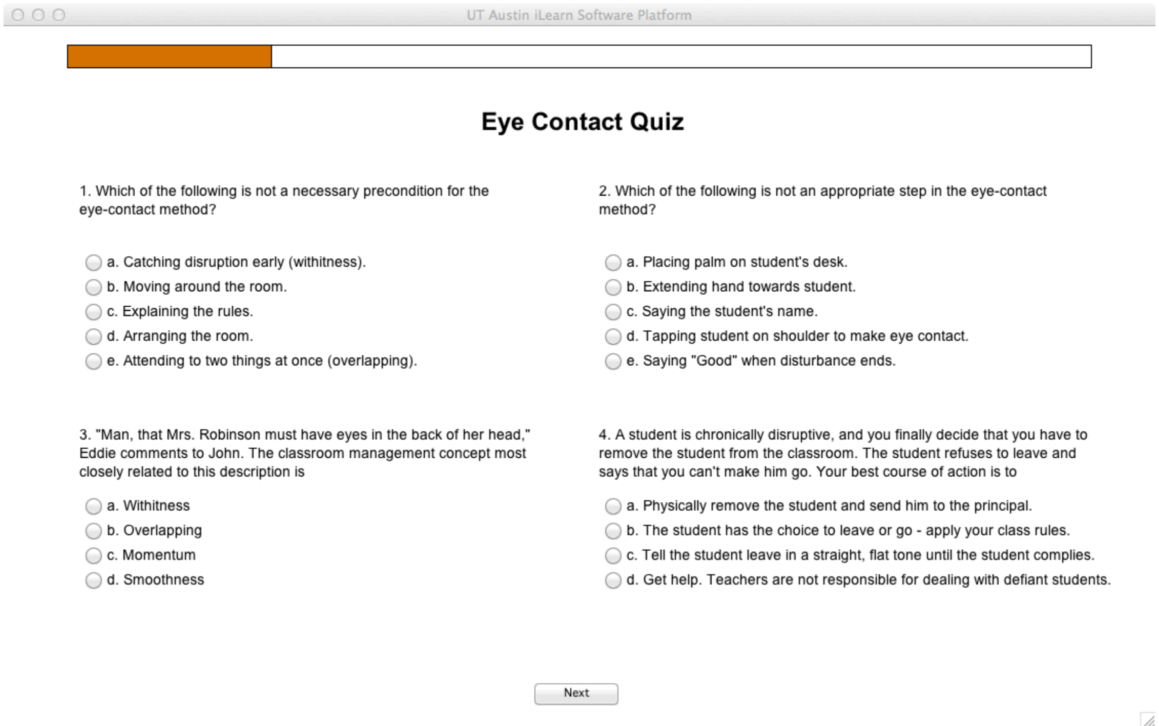
The main screen of the software, with video screen in the top left, Instant Messages panel in the top right, optional Notes window bottom right. The Start Lecture button in the bottom left begins the video.

Figure 3



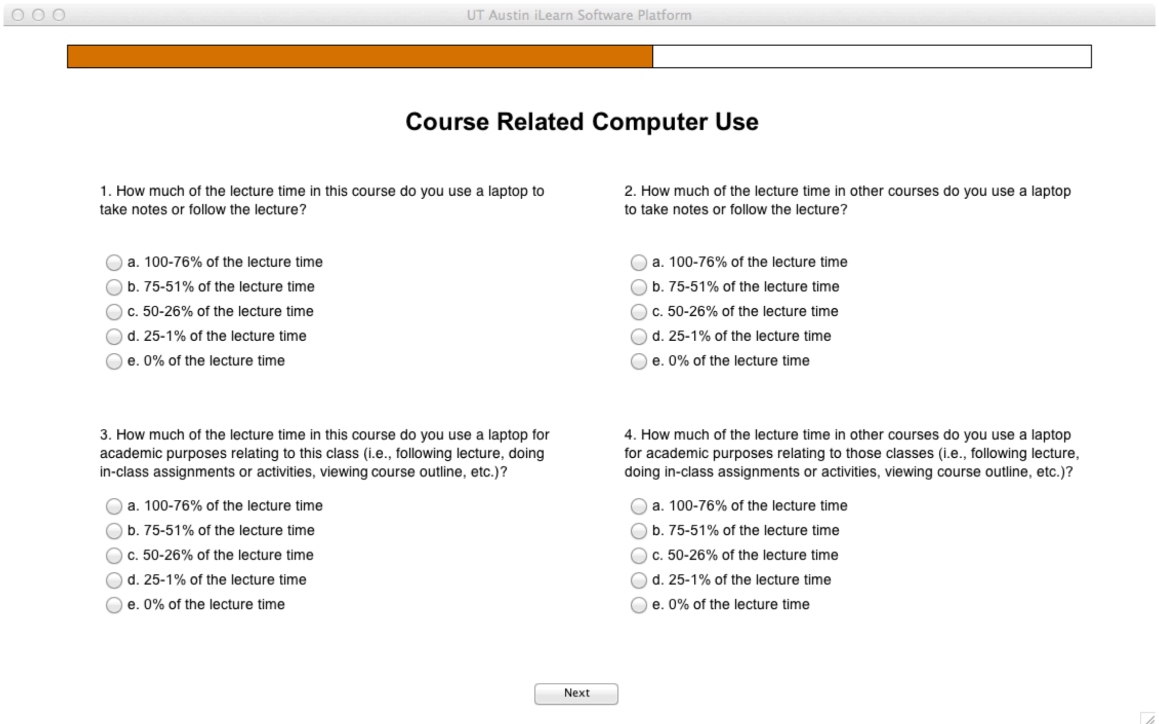
The final screen of the experiment.

Figure 4



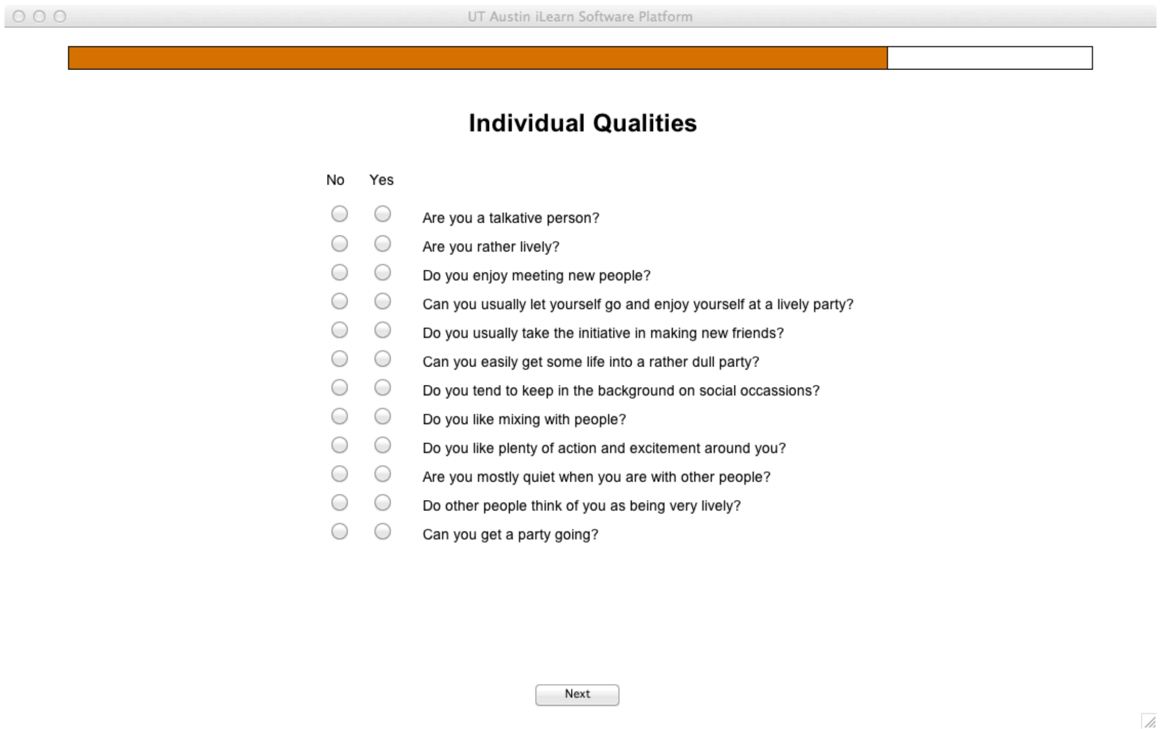
The first screen of the recall quiz.

Figure 5



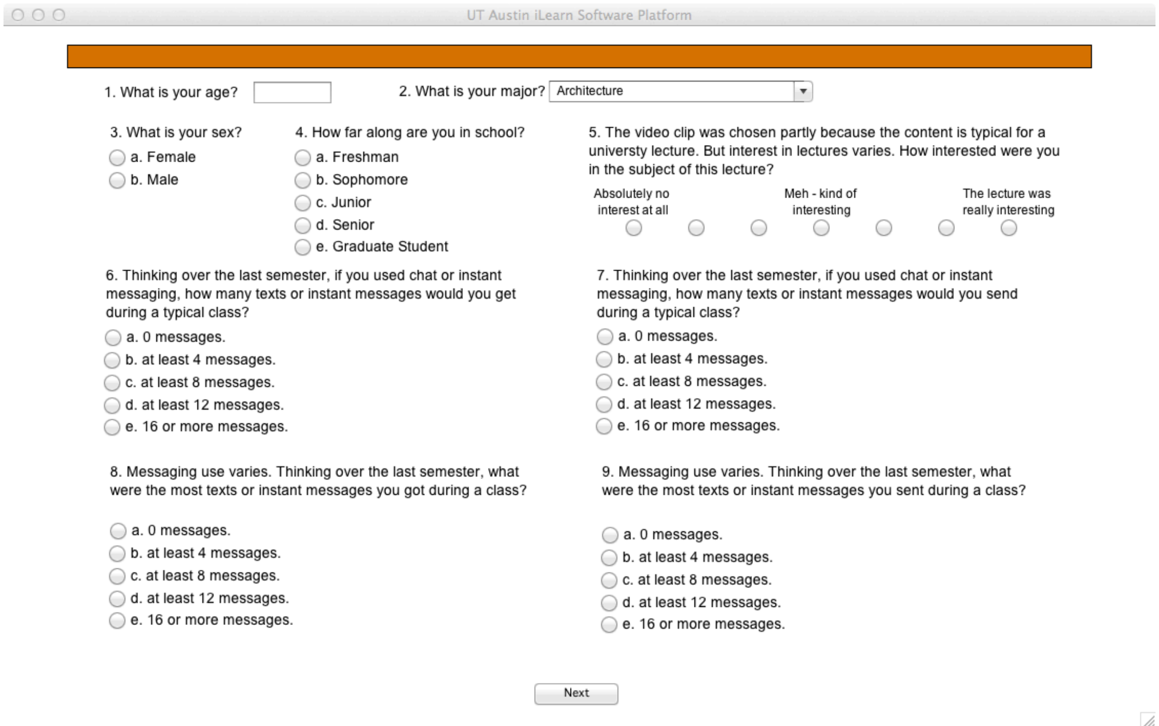
The first screen of the LES scale.

Figure 6



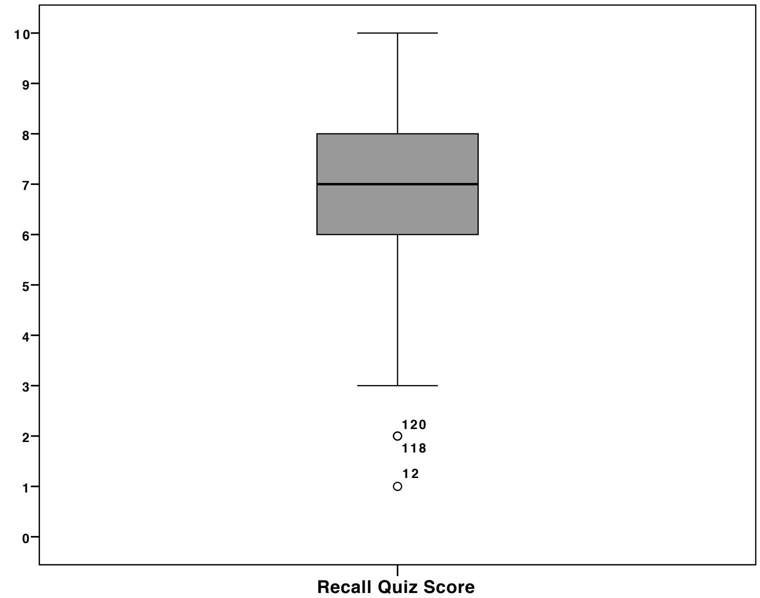
The EPQ–BV sub–scale used to assess cognitive arousal in the Concluding study.

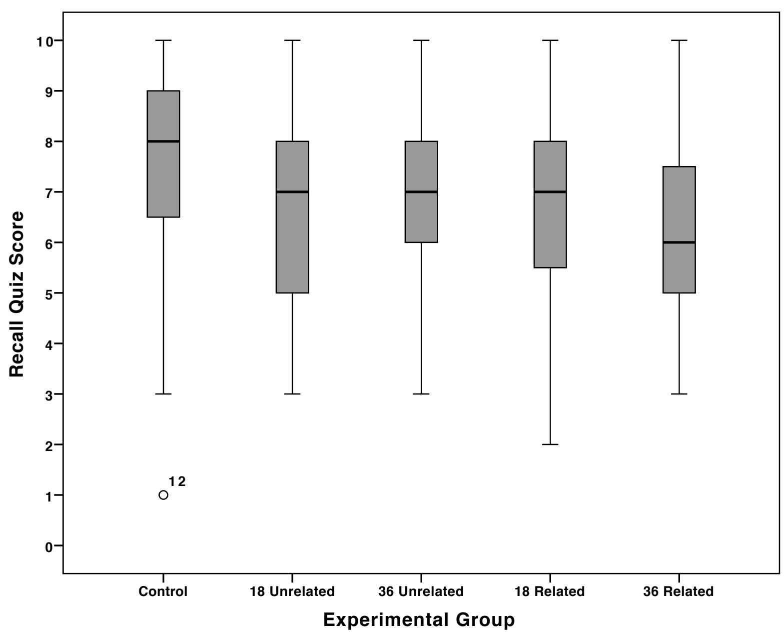
Figure 7



Demographic screen from the Concluding study.

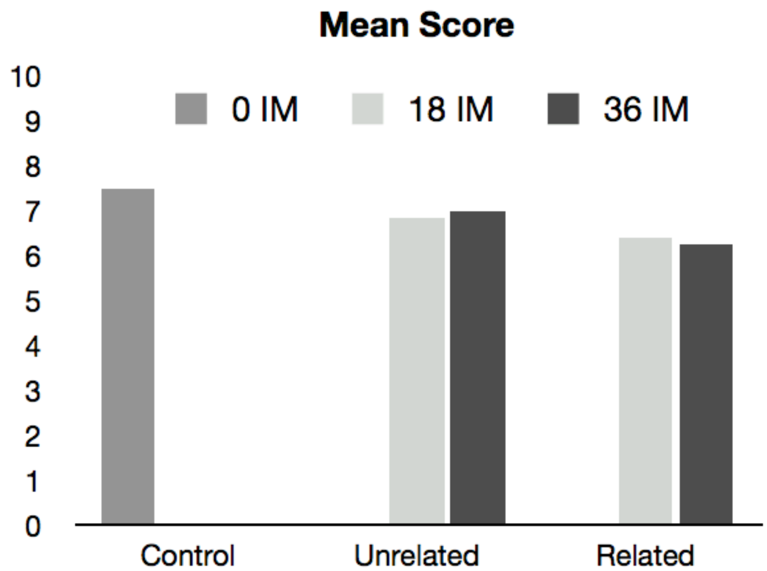
Figure 8





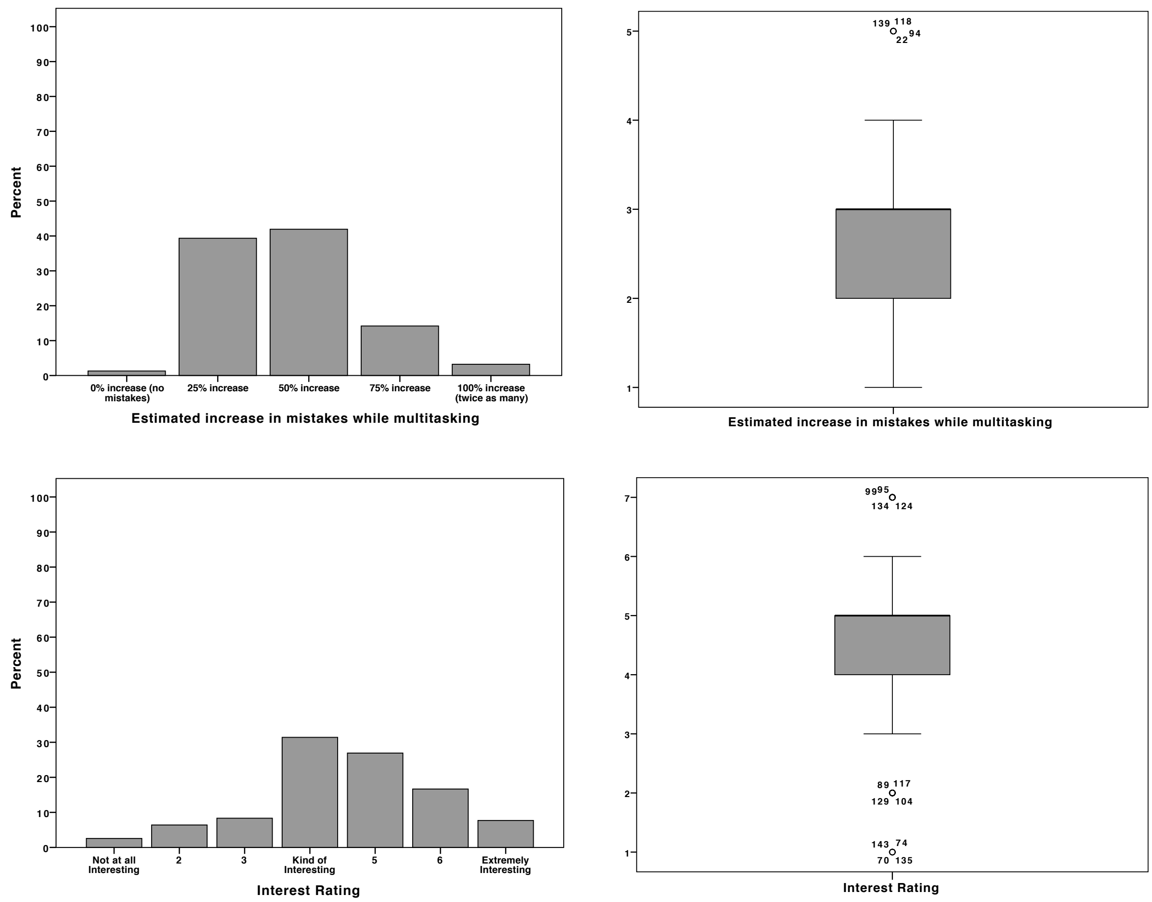
Pilot study outlying scores by sample and group.

Figure 9



Pilot study quiz scores by group.

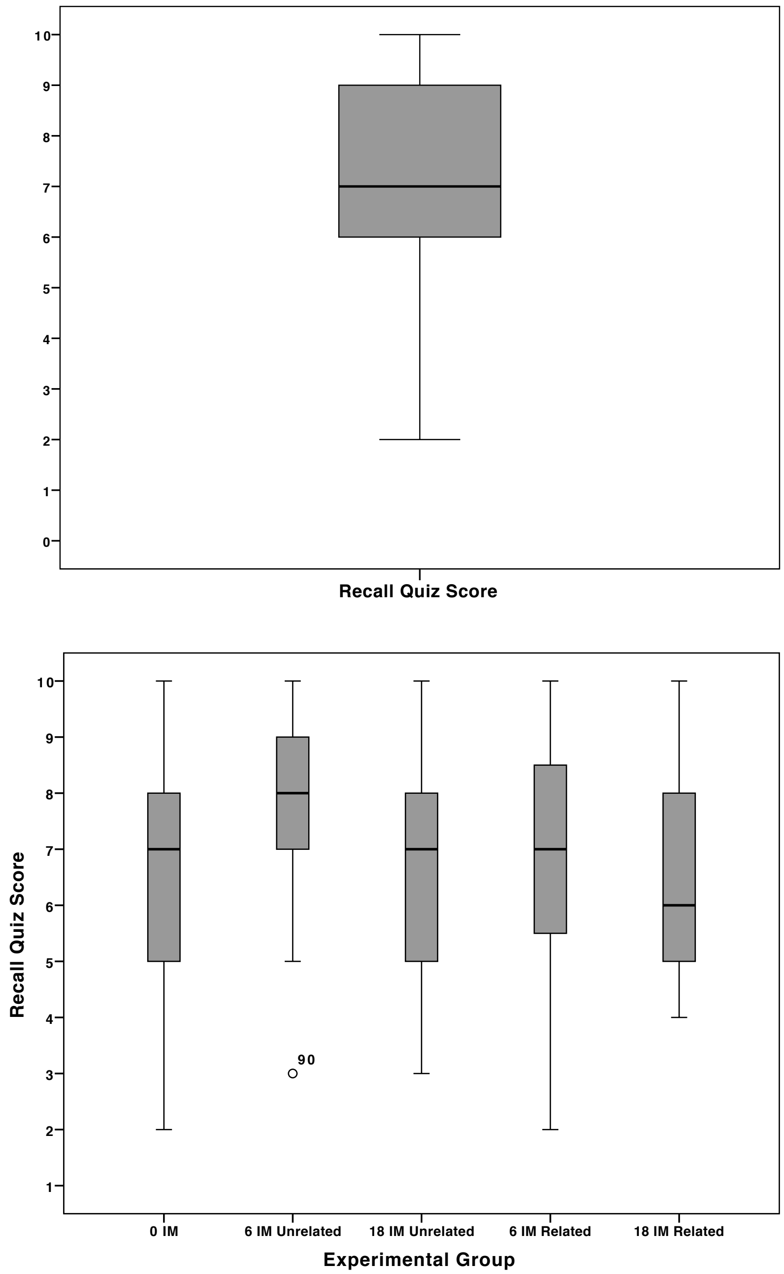
Figure 10



Pilot study distributions for interest in lecture and multitasking ability estimate.

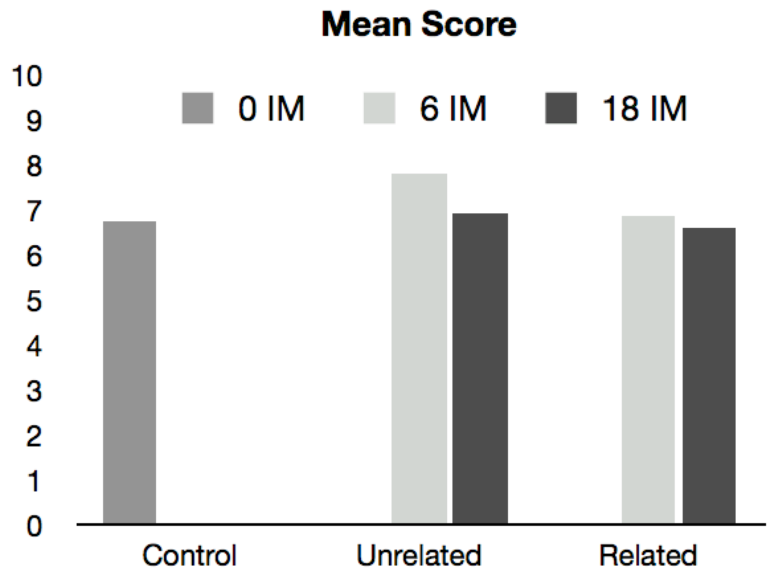
Note that one student failed to answer the multitasking question.

Figure 11



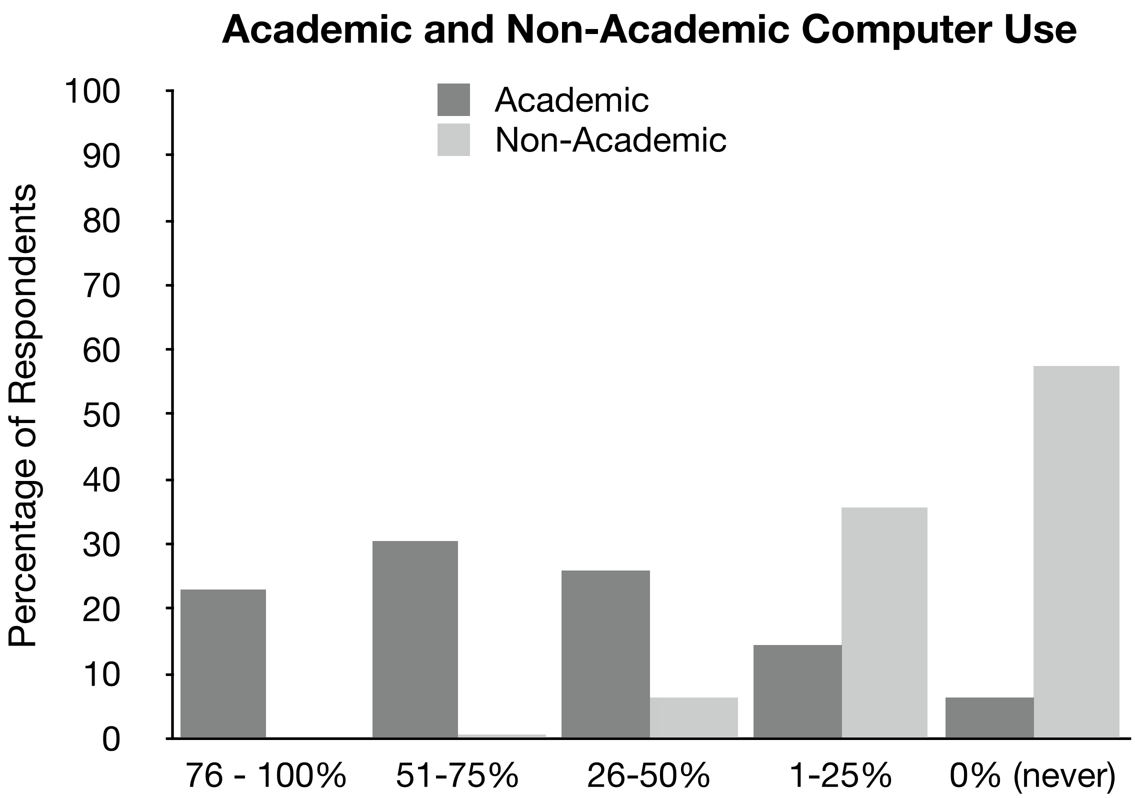
Preliminary study outlying scores by sample and group.

Figure 12



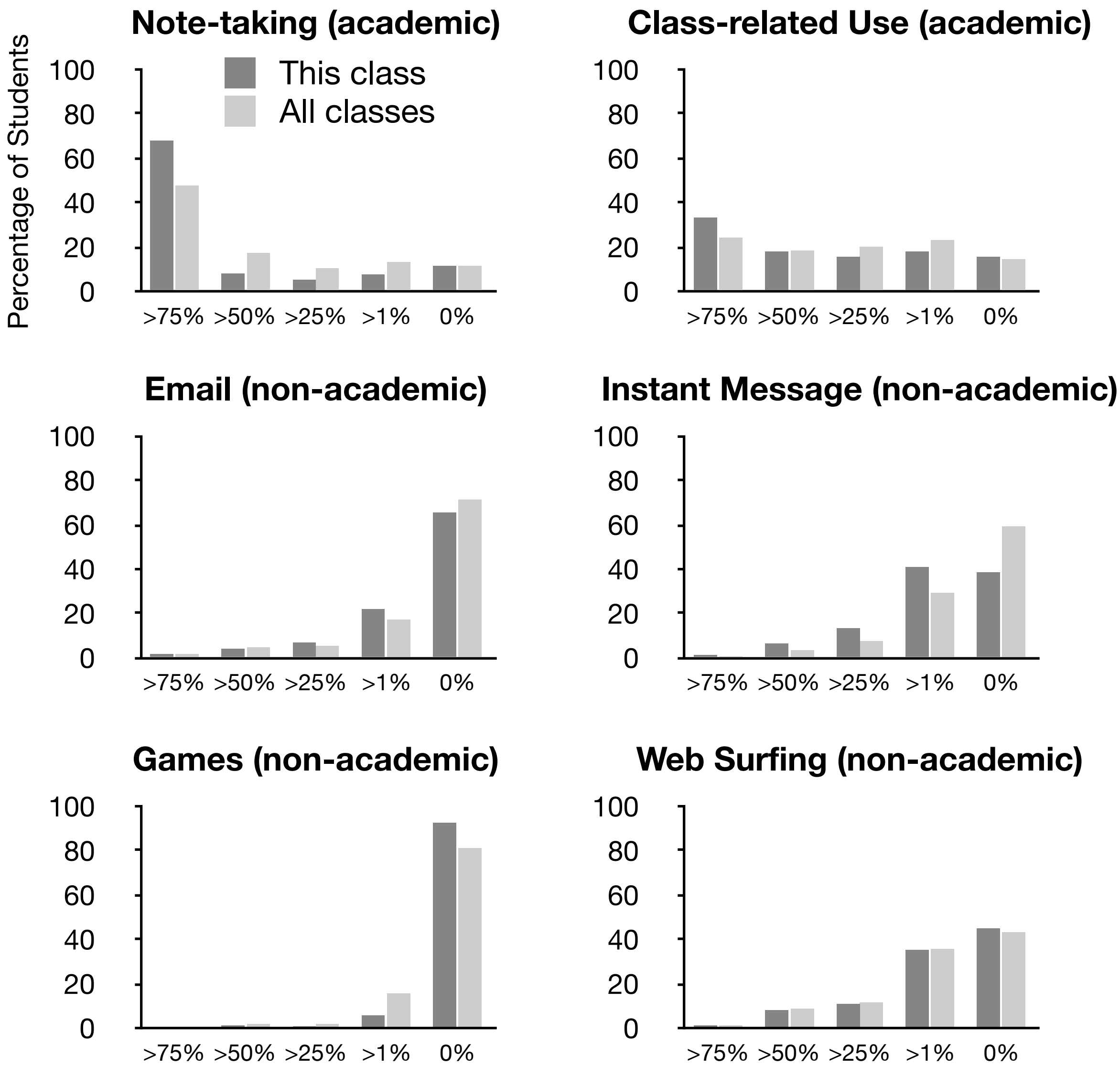
Preliminary study mean scores by experimental group.

Figure 13



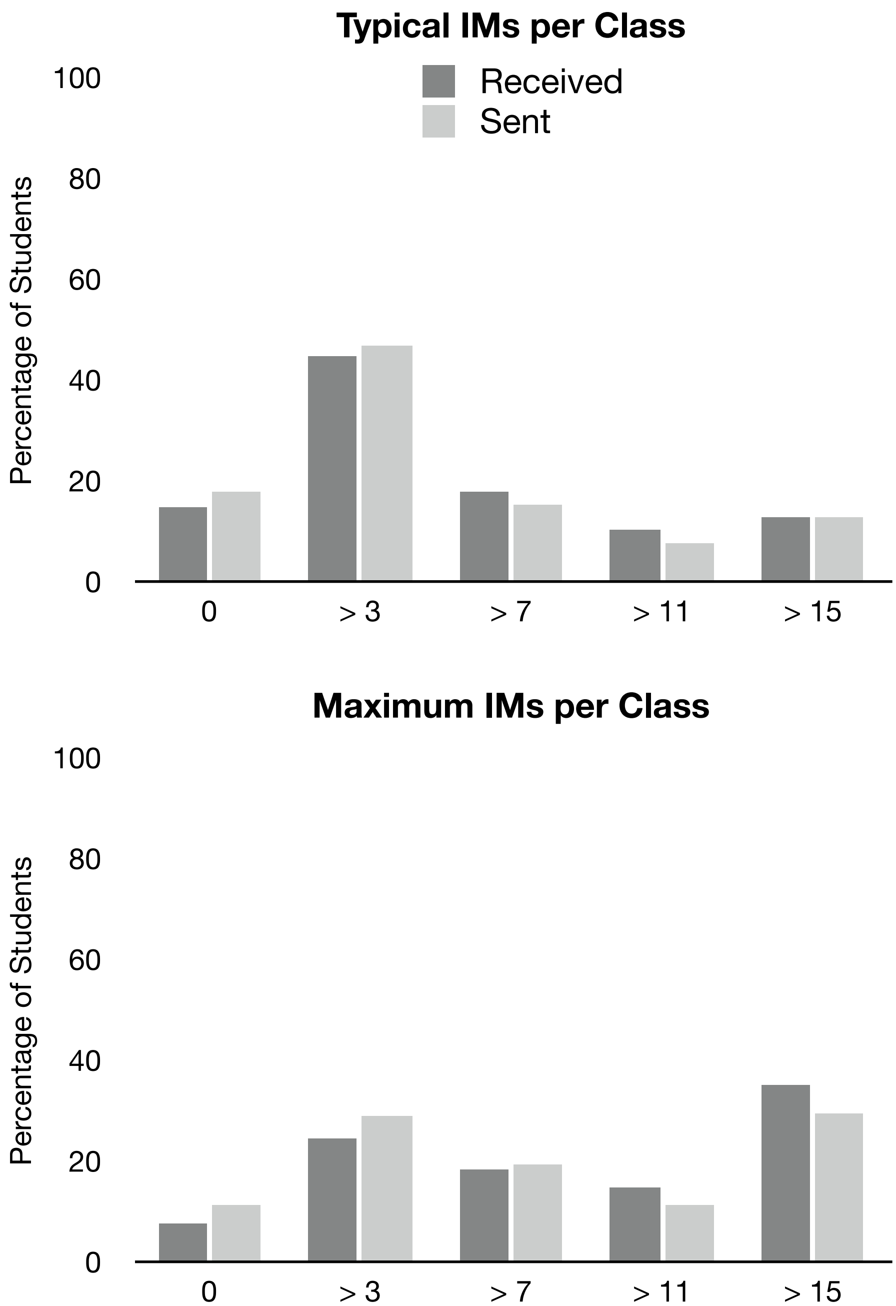
Academic versus non–academic computer use.

Figure 14



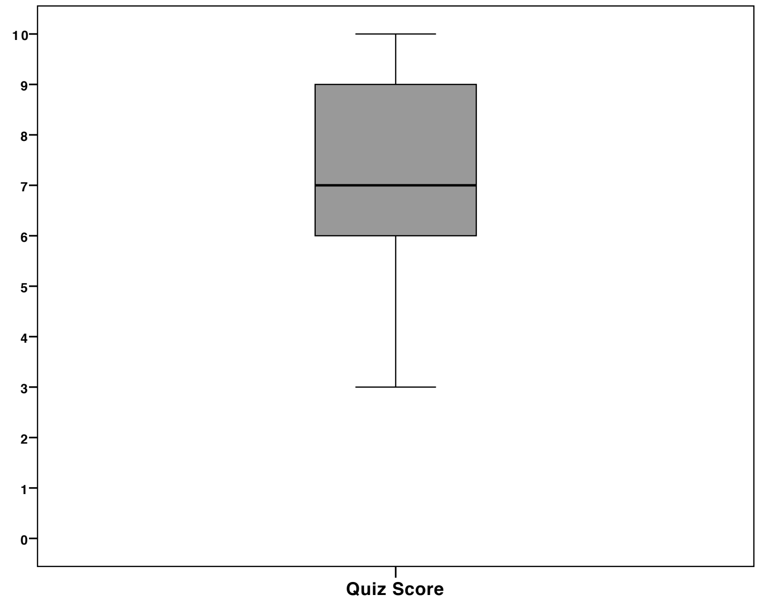
Breakdown of classroom computer use by activity.

Figure 15



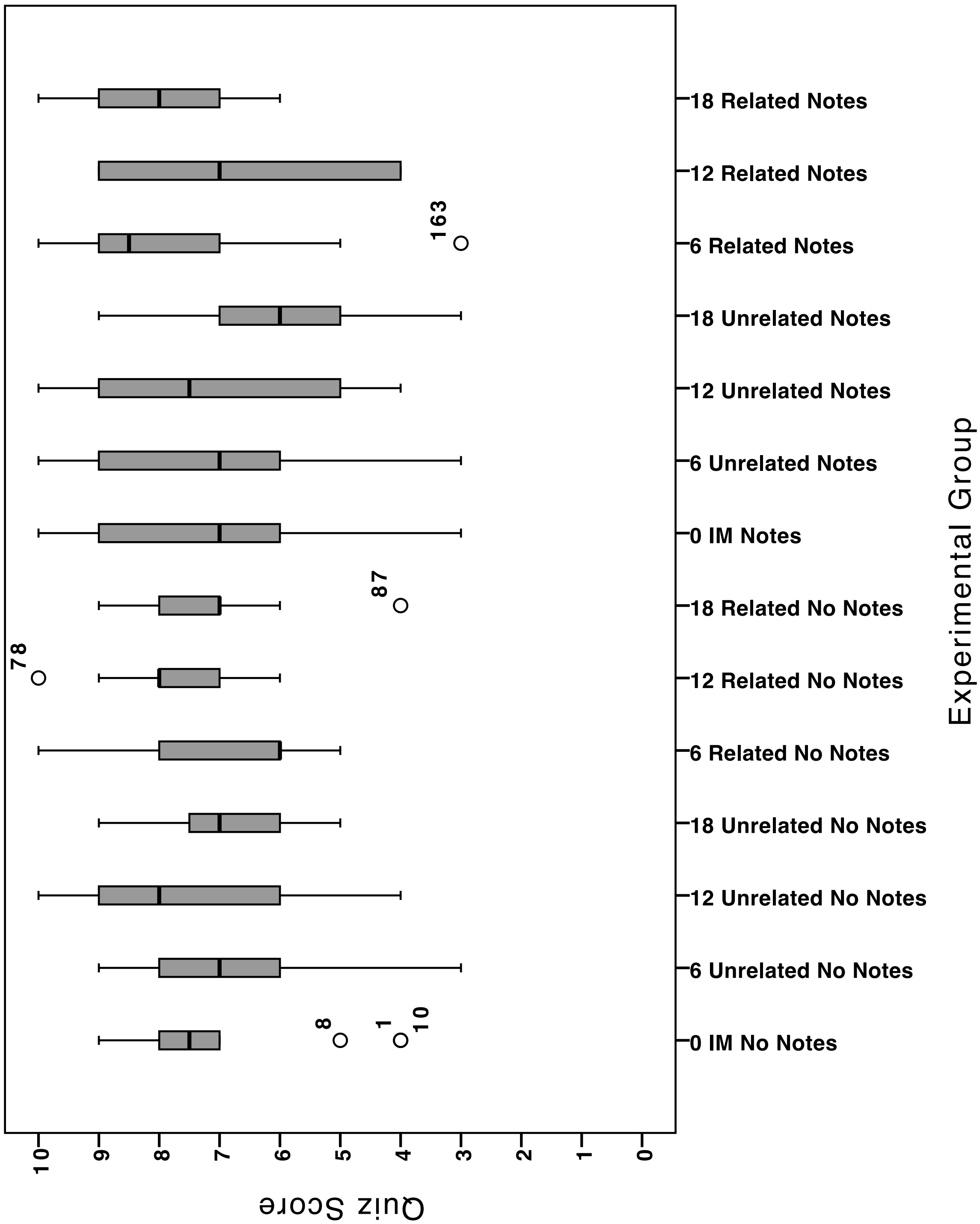
Preliminary study self–estimate for typical and maximum instant messages exchanged.

Figure 16



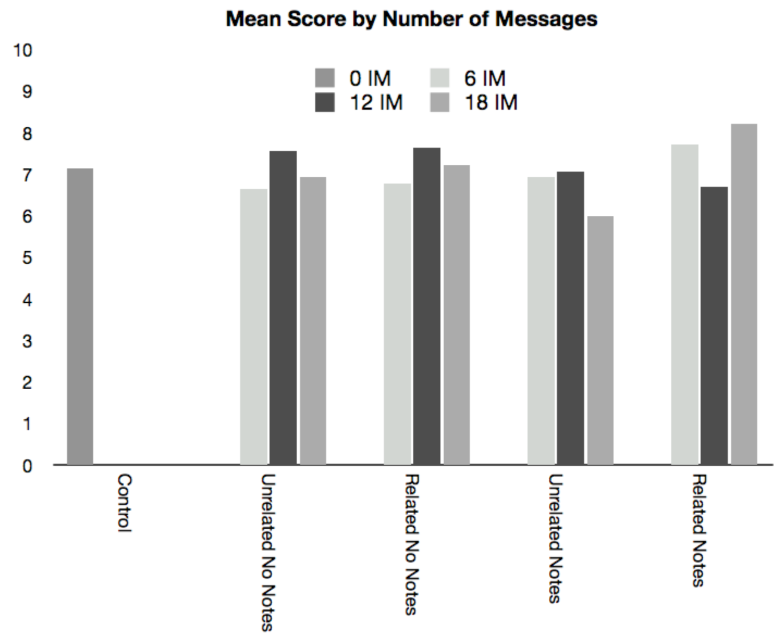
Concluding study outlying scores by sample.

Figure 17



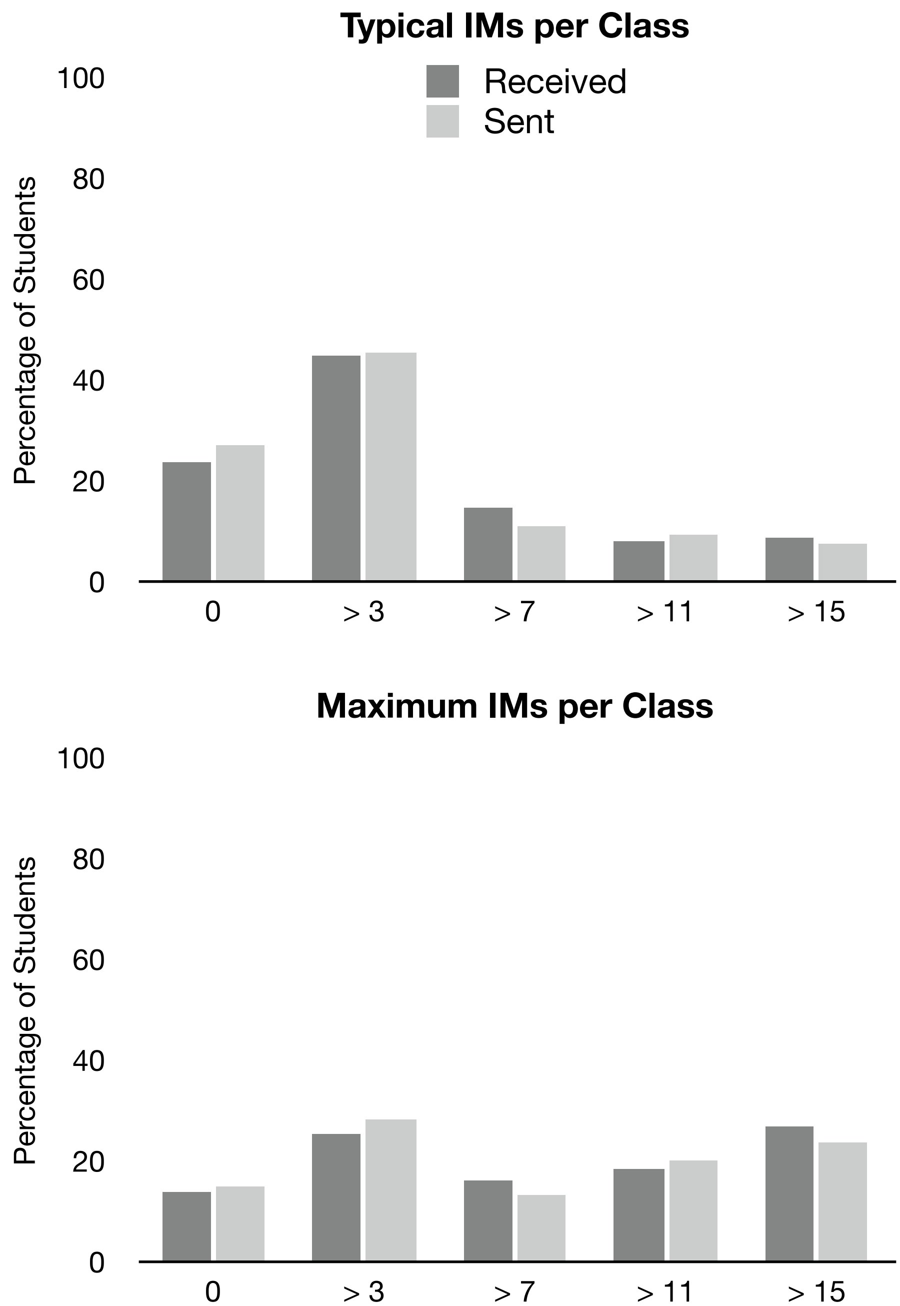
Concluding study outlying scores by group.

Figure 18



Concluding study mean scores by experimental group.

Figure 19



Concluding study self–estimate for typical and maximum instant messages exchanged.

Bibliography

Alderman, N. (1996). Central executive deficit and response to operant conditioning methods. *Neuropsychological Rehabilitation*, *6*(1), 161–186.

Altmann, E., & Trafton, J. (2002). Memory for goals: An activation–based model. *Cognitive Science*, *26*(1), 39–83.

Altmann, E., & Trafton, J. (2004). Task interruption: Resumption lag and the role of cues. In *Proceedings of the Twentieth–Sixth Annual Conference of the Cognitive Science Society*. Presented at the Proceedings of the Twentieth–Sixth Annual Conference of the Cognitive Science Society.

Anderson, D. R., Levin, S. R., & Lorch, E. P. (1977). The effects of TV program pacing on the behavior of preschool children. *AV Communication Review*, *25*(2), 159–166.

Anderson, K. J., & Revelle, W. (1982). Impulsivity, caffeine, and proofreading: A test of the Easterbrook hypothesis. *Journal of Experimental Psychology: Human Perception and Performance*, *8*(4), 614–624.

Armstrong, G. B., & Greenberg, B. S. (1990). Background television as an inhibitor of cognitive processing. *Human Communication Research*, *16*(3), 355–386.

Atkinson, R. C., & Shiffrin, R. M. (1968). Human memory: A proposed system and its control processes. In K. W. Spence (Ed.), *The Psychology of Learning and Motivation: Advances in Research and Theory* (Vol. 2, pp. 89–195). New York, NY: Academic Press.

Baddeley, A. D. (1986). *Working Memory*. Oxford, UK: Oxford University Press.

Baddeley, A. D. (2000). The episodic buffer: a new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417–423.

Baddeley, A. D. (2007). *Working Memory, Thought, and Action*. Oxford, UK: Oxford University Press.

Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. A. Bower (Ed.), *Recent Advances in Learning and Motivation* (Vol. 8, pp. 47–89). New York, NY: Academic Press.

Baddeley, A. D., & Logie, R. H. (1999). Working memory: The multiple component model. In A. Miyake & P. Shah (Eds.), *Models of Working Memory: Mechanisms of Active Maintenance and Executive Control* (pp. 28–61). Cambridge, UK: Cambridge University Press.

Baddeley, A. D., Sala, Della, S., Papagono, C., & Spinnler, H. (1997). Dual task performance in dysexecutive and non–dysexecutive patients with a frontal lesion. *Neuropsychology*, *11*(2), 187–194.

Baker, L., & Lombardi, B. R. (1985). Students' lecture notes and their relation to test performance. *Teaching of Psychology*, *12*(1), 28–32.

Bar–Ilan, J. (2008). Which h–index?—A comparison of WoS, Scopus and Google Scholar. *Scientometrics*, *74*(2), 257–271.

Barrett, L. F. (1998). Discrete emotions or dimensions? The role of valence focus and arousal focus. *Cognition & Emotion*, *12*(4), 579–599.

Barrett, L. F., & Russell, J. A. (1999). The structure of current affect: Controversies and emerging consensus. *Current Directions in Psychological Science*, *8*(1), 10–14.

Battestini, A., Setlur, V., & Sohn, T. (2010). A large scale study of text–messaging use. *Proceedings of the 12th international conference on human computer interaction with mobile devices and services*, 229–238.

Bennett, S., Maton, K., & Kervin, L. (2008). The "digital natives" debate: A critical review of the evidence. *British Journal of Educational Technology*, *39*(5), 775–786.

Borbone, M. (2009). *From Doodle to Google: Exacerbating old problems or causing new ones? The relationship between laptop usage and in–class student behavior.* (D. Barreau, Ed.) (pp. 1–93). University of North Carolina at Chapel Hill.

Bowman, L. L., Levine, L., Waite, B., & Gendron, M. (2010). Can students really multitask? An experimental study of instant messaging while reading. *Computers & Education*, *54*(4), 927–931.

Bradshaw, T., & Nichols, B. (2004). Reading At Risk: A Survey of Literary Reading in America. In D. Ball (Ed.), *Research Division Report# 46*. (pp. 1–47). Washington, DC: National Endowment for the Arts.

Broadbent, D. E. (1965). A reformulation of the Yerkes–Dodson law. *British Journal of Mathematical and Statistical Psychology*, *18*(2), 145–157.

Broadbent, D. E. (1971). *Decision and Stress*. New York, NY: Academic Press.

Brooks, D. (2001, April 30). Time to do everything except think. *Newsweek*, *71*.

Brown, J. (2000). Growing up digital: How the web changes work, education, and the ways people learn. *Change: The Magazine of Higher Learning*, *32*(2), 11–20.

Bugeja, M. J. (2006). Facing the Facebook. *The Chronicle of Higher Education*, *52*(21), 1.

Bugeja, M. J. (2007a). Distractions in the wireless classroom. *The Chronicle of Higher Education*, *53*(21), C1.

Bugeja, M. J. (2007b). Second thoughts about Second Life. *The Chronicle of Higher Education*, *54*(3).

Bugeja, M. J. (2008). Classroom clickers and the cost of technology. *Chronicle of Higher Education*, A31–A32.

Bullen, M., Morgan, T., & Qayyum, A. (2010). Digital learners in higher education – Generation is not the issue. *Canadian Journal of Learning and Technology*, *37*(1), 1–24.

Caldwell, J. (2007). Clickers in the large classroom: Current research and best–practice tips. *CBE Life Sciences Education*, *6*(1), 9–20.

Campbell, J. B., & Hawley, C. W. (1982). Study habits and Eysenck's theory of Extraversion–Introversion. *Journal of Research in Personality*, *16*(2), 139–146.

Carrier, L. M., Cheever, N., Rosen, L., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of Americans. *Computers in Human Behavior*, *25*(2), 483–489.

Carroll, J. M., Yik, M. S. M., Russell, J., & Barrett, L. F. (2004). On the psychometric principles of affect. *Review of General Psychology*, *3*(1), 14–22.

Caruso, J., & Salaway, G. (2007). *The ECAR Study of Undergraduate Students and Information Technology, 2007: Key Findings*. Educause Center for Applied Research.

Cisco Systems, Inc. (2012) *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2011–2016*. Downloaded February 16, 2012 from: https://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

Cohen, M. X., Young, J., Baek, J. M., Kessler, C., & Ranganath, C. (2005). Individual differences in extraversion and dopamine genetics predict neural reward responses. *Cognitive Brain Research*, *25*(3), 851–861.

Connor, E. (2009). Perceptions and uses of clicker technology. *Journal of Electronic Resources in Medical Libraries*, *6*(1), 19–32.

Cowan, N. (2001). The magical number 4 in short–term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, *24*(1), 87–114.

Craik, F. I. M., Govoni, R., Naveh–Benjamin, M., & Anderson, N. D. (1996). The effects of divided attention on encoding and retrieval processes in human memory. *Journal of Experimental Psychology: General*, *125*(2), 159–180.

Diamanduros, T., Jenkins, S., & Downs, E. (2007). Analysis of technology ownership and selective use among undergraduates. *College Student Journal*, *41*(4), 970–976.

Dzubak, C. (2008). Multitasking: The good, the bad, and the unknown. *Synergy*, *1*(2), 1–12.

Easterbrook, J. A. (1959). The effect of emotion on cue utilization and the organization of behavior. *Psychological Review*, *66*(3), 183–201. American Psychological Association.

Entner, R. (2010, March 26). *Smartphones to Overtake Feature Phones in U.S. by 2011*. Retrieved July 14, 2011, from The Nielsen Company Web site: http://blog.nielsen.com/nielsenwire/consumer/smartphones-to-overtake-feature-phones-in-u-s-by-2011.

Eysenck, H. J. (1967). *The Biological Basis of Personality*. Springfield, IL: Thomas Publishing.

Eysenck, H. J., & Eysenck, S. B. G. (1994). *Manual of the Eysenck Personality Questionnaire*. San Diego, CA: Educational and Industrial Testing Service.

Eysenck, M. W. (1976). Arousal, learning, and memory. *Psychological Bulletin*, *83*(3), 389–404.

Eysenck, M. W. (1981). *Attention and Arousal: Cognitive Performance* (p. 209). New York, NY: Springer–Verlag.

Eysenck, M. W., & Calvo, M. G. (1992). Anxiety and performance: The processing efficiency theory. *Cognition & Emotion*, *6*(6), 409–434.

Eysenck, M. W., Derakshan, N., Santos, R., & Calvo, M. G. (2007). Anxiety and cognitive performance: Attentional control theory. *Emotion*, *7*(2), 336–353.

Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the psychoticism scale. *Personality and Individual Differences*, *6*(1), 21–29.

Felton, D. F. (2007). Cognitive load and classroom teaching: The double–edged sword of automaticity. *Educational Psychologist*, *42*(3), 123–137.

Flanagin, A. J. (2005). IM online: Instant messaging use among college students. *Communication Research Reports*, *22*(3), 175–187.

Foster, A. L. (2009). Law professors rule laptops out of order in class. *The Chronicle of Higher Education*, *54*(40), A1.

Fox, A., Rosen, J., & Crawford, M. (2009). Distractions, distractions: Does instant messaging affect college students' performance on a concurrent reading comprehension task? *CyberPsychology & Behavior*, *12*(1), 51–53.

Francis, L. J., Brown, L. B., & Philipchalk, R. (1992). The development of an abbreviated form of the Revised Eysenck Personality Questionnaire (EPQR–A): Its use among students in England, Canada, the USA and Australia. *Personality and Individual Differences*, *13*(4), 443–449.

Frand, J. (2000). The information age mindset. *Educause Review*, *35*(5), 15–24.

Freeburne, C. M., & Fleischer, M. S. (1952). The effect of music distraction upon reading rate and comprehension. *Journal of Educational Psychology*, *43*(2), 101–109. Warwick & York.

Fried, C. (2008). In–class laptop use and its effects on student learning. *Computers & Education*, *50*(3), 906–914.

Furnham, A., & Allass, K. (1999). The influence of musical distraction of varying complexity on the cognitive performance of extroverts and introverts. *European Journal of Personality*, *13*(1), 27–38.

Furnham, A., & Bradley, A. (1997). Music while you work: The differential distraction of background music on the cognitive test performance of introverts and extraverts. *Applied Cognitive Psychology*, *11*(5), 445–455.

Furnham, A., Gunter, B., & Peterson, E. (1994). Television distraction and the performance of introverts and extroverts. *Applied Cognitive Psychology*, *8*(7), 705–711.

Gale, A. (1973). The psychophysiology of individual differences: Studies of extraversion and the EEG. In P. Kline (Ed.), *New Approaches in Psychological Measurement*. London, UK: Wiley

Geen, K. G. (1976). *Personality, the Skein of Behavior*. St. Louis, MO.

Giles, J. (2005). Internet encyclopaedias go head to head. *Nature*, *438*(7070), 900–901.

Gonder, J. (2010). *Text Messaging in Class: Extraversion as a Mediating Factor in Distraction*. Unpublished Master's Thesis, California State University, Chico, Chico, California.

Griswold, W., McDonnell, T., & Wright, N. (2005). Reading and the reading class in the Twenty–First century. *Annual Review of Sociology*, *31*(1), 127–141.

Guerrero, C. (2010). *Fostering Active Learning Through the Use of Feedback Technologies and Collaborative Activities in a Postsecondary Setting*. Unpublished Doctoral Dissertation, University of Texas at Austin.

Hadwin, A. F., Kirby, J. R., & Woodhouse, R. A. (1999). Individual differences in notetaking, summarization, and learning from lectures. *Alberta Journal of Educational Research*, *45*(1), 1–17.

Hammer, R., Ronen, M., Sharon, A., Lankry, T., Huberman, Y., & Zamtsov, V. (2010). Mobile culture in college lectures: Instructors' and students' perspectives. *Interdisciplinary Journal of E–Learning and Learning Objects*, *6*, 293–304.

Hartman, A., Pickering, R. M., & Wilson, B. A. (1992). Is there a central executive deficit after severe head injury? *Clinical Rehabilitation*, *6*(1), 133–140.

Helsper, E. J., & Eynon, R. (2010). Digital natives: Where is the evidence? *British Educational Research Journal*, *36*(3), 503–520.

Hembrooke, H., & Gay, G. (2003). The laptop and the lecture: The effects of multitasking in learning environments. *Journal of Computing in Higher Education*, *15*(1), 46–64.

Hill, A. (1975). Extraversion and variety−seeking in a monotonous task. *British Journal of Psychology*, *66*(1), 9–13.

Hornik, R. (1981). Out–of–school television and schooling: Hypotheses and methods. *Review of Educational Research*, *51*(2), 193–214.

Howarth, E., & Eysenck, H. J. (1968). Extraversion, arousal, and paired–associate recall. *Journal of Experimental Research in Personality*, *3*(2), 114–116.

Hu, M., Lim, E., Sun, A., Lauw, H., & Vuong, B. (2007). Measuring article quality in Wikipedia: models and evaluation. *Proceedings of the Sixteenth ACM Conference on Conference on information and Knowledge Management*, 243–252.

Jacso, P. (2005). As we may search–comparison of major features of the Web of Science, Scopus, and Google Scholar citation–based and citation–enhanced databases. *Current science*, *89*(9), 1537–1547.

Johnson, D. L., Wiebe, J. S., Gold, S. M., Andreasen, N. C., Hichwa, R. D., Watkins, G. L., & Boles Ponto, L. L. (1999). Cerebral blood flow and personality: A positron emission tomography study. *American Journal of Psychiatry*, *156*(2), 252–257. Am Psychiatric Assoc.

Jones, F. H. (1978). The eye–contact method. *Instructor*, *88*(4), 64.

Kahneman, D. A. (1973). Attention and Effort. *Englewood Cliffs, NJ: Prentice Hall*.

Katayama, A. D., Shambaugh, R. N., & Doctor, T. (2010). Promoting knowledge Transfer with electronic note taking. *Teaching of Psychology*, *32*(2), 129–131.

Kay, R. H., & Lauricella, S. (2011a). Gender differences in the use of laptops in higher education: A formative analysis. *Journal of Educational Computing Research*, *44*(3), 361–380.

Kay, R. H., & Lauricella, S. (2011b). Exploring the benefits and challenges of using laptop computers in higher education classrooms: A formative analysis. *Canadian Journal of Learning and Technology/La revue canadienne de l'apprentissage et de la technologie*, *37*(1), 2–18.

Kay, R. H., & Lauricella, S. (2011c). Unstructured vs. structured use of laptops in higher education. *Journal of Information Technology Education*, *10*(1), 33–42.

Kinzie, M., Whitaker, S., & Hofer, M. (2005). Instructional uses of instant messaging (IM) during classroom lectures. *Educational Technology & Society*, *8*(2), 150–160.

Kirschner, P. A., & Karpinski, A. (2010). Facebook® and academic performance. *Computers in Human Behavior*, *26*(6), 1237–1245.

Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one's own incompetence lead to inflated self–assessments. *Journal of Personality and Social Psychology*, *77*(6), 1121–1134.

Lauricella, S., & Kay, R. H. (2010). Assessing laptop use in higher education classrooms: The laptop effectiveness scale (LES). *Australasian Journal of Educational Technology*, *26*(2), 151–163.

Leach, J., Lewin, C., & Pearson, M. (2007). Technological, pedagogical and cultural shifts. *Learning, Media, & Technology*, *32*(1), 1–4.

Lenhart, A., Rainie, L., & Lewis, O. (2001). *Teenage Life Online*. *The Rise of the Instant Message Generation and the Internet's Impact on Friendships and Family Relationship*s (pp. 1–46). Online: www. pewinternet. org.

Lente, H. V., & Rip, A. (1998). *Expectations in technological developments: An example of prospective structures to be filled in by agency*. In C. Disco & B. van de Meulen (Eds.), *Getting New Technologies Together. Studies in Making Sociotechnical Order* (pp. 203–231). Berlin, FDR: Walter de Gruyter.

Leu, D. J., Jr., Zawilinski, L., Castek, J., Banerjee, M., Housand, B., Liu, Y., & O'Neil, M. (2007). What is new about the new literacies of online reading comprehension? In L. S. Rush, A. J. Eakle, & A. Berger (Eds.), *Secondary School Reading and Writing: What Research Reveals for Classroom Practices* (pp. 37–68). Chicago, IL: National Council of Teachers of English.

Levy, D. M., & Nardick, D. L. (2011, May 8). No cellphone? No internet? So much less stress. *The Chronicle of Higher Education*. Retrieved May 30, 2011 from http://chronicle.com/article/No-Cellphone-No-Internet-So/127391/

Levy, S. (2006, March 27). Some attention must be paid! *Newsweek*.

Liu, Z. (2005). Reading behavior in the digital environment: Changes in reading behavior over the past ten years. *Journal of Documentation*, *61*(6), 700–712.

Logie, R. H., Cocchini, G., Delia Sala, S., & Baddeley, A. D. (2004). Is there a specific executive capacity for dual task coordination? Evidence from Alzheimer's disease. *Neuropsychology*, *18*(3), 504–513.

Mackworth, J. F. (1969). *Vigilance and Habituation*. Harmondsworth: Penguin Books.

Mahometa, M. J., Eakin, T., & Smith, S. (2008). *Analysis of ITS Networking Public Network Access Usage*. Retrieved September 16, 2010 from The

University of Texas at Austin, Information Technology Services Web site: https://www.utexas.edu/its/network/reports/PNASpring2008UsageReport.pdf.

Malmo, R. B. (1957). Anxiety and behavioral arousal. *Psychological Review*, *64*(5), 276–287.

Mangan, K. S. (2001). Cutting the power. *The Chronicle of Higher Education*, *48*(2), A43–A44.

Marois, R., & Ivanoff, J. (2005). Capacity limits of information processing in the brain. *Trends in Cognitive Sciences*, *9*(6), 296–305.

Maxwell, N. G. (2007). From Facebook to Folsom Prison blues: How banning laptops in the classroom made me a better law school teacher. *Richmond Journal of Law and Technology*, *14*(2), 1–43.

McCreary, J. (2009). The laptop–free zone. *Valparaiso University Law Review*, *43*(3), 989–1044.

McGehee, W., & Gardner, J. E. (1949). Music in a complex industrial job. *Personnel Psychology*, *2*, 405–417.

McWilliams, G. (2005, October 15). The laptop backlash. *The Wall Street Journal*, p. B1.

Meade, P. (2003, January 20). Multitasking helps you to stay ahead. *The San Diego Union–Tribune*, p. C3.

Meyer, D., & Kieras, D. (1997). A computational theory of executive cognitive processes and multiple–task performance: Part 1. Basic mechanisms. *Psychological Review*, *104*(1), 3.

Miles, J., & Shevlin, M. (2001). *Applying Regression & Correlation: A Guide for Students and Researchers*. London, UK: Sage Publications.

Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81–97.

Mokhtari, K., Reichard, C. A., & Gardner, A. (2009). The impact of internet and television use on the reading habits and practices of college students. *Journal of Adolescent & Adult Literacy*, *52*(7), 609–619.

Moore, G. E. (1965). Cramming more components onto integrated circuits. *Electronics*, *38*(8), 1–4.

Morgan, M. (1980). Television viewing and reading: Does more equal better? *Journal of Communication*, *30*(1), 159–165.

Nairne, J. (2002). Remembering over the short–term: The case against the standard model. *Annual Review of Psychology*, *53*(1), 53–81.

Naveh–Benjamin, M., Guez, J., & Marom, M. (2003). The effects of divided attention at encoding on item and associative memory. *Memory & Cognition*, *31*(7), 1021–1035.

Naveh–Benjamin, M., Kilb, A., & Fisher, T. (2006). Concurrent task effects on memory encoding and retrieval: Further support for an asymmetry. *Memory & Cognition*, *34*(1), 90–101.

Negroponte, N. (1995). *Being Digital* (p. 243). New York, NY: Knopf.

Netcraft Ltd. (2010). *Web Server Survey*. Retrieved September 15, 2010, from http://news.netcraft.com/archives/category/web-server-survey/

Newman, R. I., Jr., Hunt, D. L., & Rhodes, F. (1966). Effects of music on employee attitude and productivity in a skateboard factory. *Journal of Applied Psychology*, *50*(6), 493–496.

Olsen, F. (2002). Duke U. decides against requiring freshmen to own laptops. *The Chronicle of Higher Education*, (Jan 11).

Ophir, E., Nass, C., & Wagner, A. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(37), 15583.

Oyzon, M., & Olmos, O. (2010). Students' notes and their relation to comprehension and recall of lecture information. *Alipato: A Journal of Basic Education*, *3*(3), 7–18.

Pashler, H., Johnston, J., & Ruthruff, E. (2001). Attention and performance. *Annual Review of Psychology*, *52*(1), 629–651.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, Design and Analysis: An Integrated Approach*. Hillsdale, NJ: Laurence Erlbaum.

Pool, M. M., van der Voort, T. H. A., Beentjes, J. W. J., & Koolstra, C. M. (2000). Background television as an inhibitor of performance on easy and difficult homework assignments. *Communication Research*, *27*(3), 293–326.

Prensky, M. (2001a). Digital natives, digital immigrants part 1. *On the Horizon*, *9*(5), 1–6.

Prensky, M. (2001b). Digital natives, digital immigrants part 2: Do they really think differently? *On the Horizon*, *9*(6), 1, 3–6.

Quade, A. M. (1996). An assessment of retention and depth of processing associated with notetaking using traditional pencil and paper and an on–line notepad during computer–delivered instruction. *Proceedings of the National Convention of the Association for Educational Communication and Technology*. 559–570.

Read, B. (2011). A law professor bans laptops from the classroom. *The Chronicle of Higher Education*, *52*(31), 1–2.

Richendoller, N. R., & Weaver, J. B., III. (1994). Exploring the links between personality and empathic response style. *Personality and Individual Differences*, *17*(3), 303–311.

Rojstaczer, S. (2010). Grading in American colleges and universities. Teachers College Record, Date Published: March 04, 2010 http://www.tcrecord.org ID Number: 15928, Date Accessed: 3/4/2010 1:46:08 PM

Ryder, J. (2000). Universal computer access for students – a classroom experiment in computer science. *Journal of Computing Sciences in Colleges*, *15*(5), 52–60.

Sato, T. (2005). The Eysenck personality questionnaire brief version: Factor structure and reliability. *The Journal of Psychology*, *139*(6), 545–552.

Shadish, W. R., Cook, T., & Campbell, D. T. (2002). *Experimental and Quasi–Experimental Designs for Generalized Causal Inference*. Belmont, CA: Wadsworth Cengage Learning.

Shiu, E., & Lenhart, A. (2004). How Americans use instant messaging. *Pew Internet and American Life Project*.

Smith, S. D., Caruso, J. B., & Kim, J. (2010). *The ECAR Study of Undergraduate Students and Information Technology, 2010*. EDUCAUSE Center for Applied Research.

Smith, S. D., Salaway, G., & Caruso, J. (2009). *The ECAR study of undergraduate students and information technology, 2009*. EDUCAUSE Center for Applied Research (Vol. 24, p. 2009).

Stelmack, R. M., & Wilson, X. G. (1982). Extraversion and the effects of frequency and intensity on the brainstem evoked response. *Personality and Individual Differences*, *3*(4), 373–380.

Tapscott, D. (1998). *Growing up Digital: The Rise of the Net Generation*. New York, NY: McGraw–Hill.

Thayer, R. E. (1967). Measurement of activation through self–report. *Psychological Reports*, *20*(2), 663–678.

Thayer, R. E. (1970). Activation states as assessed by verbal report and four psychophysiological variables. *Psychophysiology*, *7*, 86–94.

Thayer, R. E. (1978). Toward a psychological theory of multidimensional activation (arousal). *Motivation and Emotion*, *2*(1), 1–34. Springer.

Trafton, J., Altmann, E., Brock, D., & Mintz, F. (2003). Preparing to resume an interrupted task: Effects of prospective goal encoding and retrospective rehearsal. *International Journal of Human–Computer Studies*, *58*(5), 583–603.

Uhrbrock, R. (1961). Music on the job: Its influence on worker morale and productivity. *Personnel Psychology*, *14*(1), 9–38.

Veen, W., & Ben Vrakking. (2006). *Homo Zappiens: Growing up in a Digital Age* (p. 160). London, UK: Network Continuum Education.

Vingoe, F. (1966). Validity of the Eysenck extraversion scale as determined by self–ratings in normals. *British Journal of Social & Clinical Psychology*, *5*(1), 89–91.

Vingoe, F. (1968). Validity of the Eysenck extraversion scale: replication and extension. *Psychological Reports*, *22*(3), 706–708.

Vogel, D., Kennedy, D. M., & Kwok, R. (2007). Do mobile device applications lead to learning? *Conference Proceeding*, 1–10.

Wallis, C. (2010). *The impacts of media multitasking on children's learning & development: Report from a research seminar*. (pp. 1–32). New York, NY: The Joan Ganz Cooney Center at Sesame Workshop.

Wallis, C. (2006, March 27). The Multitasking Generation: they're e–mailing, IMing and downloading while writing the history essay. What is all that digital juggling doing to kids' brains and their family life? *Time*, *167*, 48–55.

Wang, M., Shen, R., Novak, D., & Pan, X. (2008). The impact of mobile learning on students' learning behaviours and performance: Report from a large blended classroom. *British Journal of Educational Technology*, 1–23.

Watkins, E., & Sabella, M. (2008). Examining the effectiveness of clickers on promoting learning by tracking the evolution of student responses. In *AIP Conference Proceedings, 1064*. 223–226.

Watson, J., & Strayer, D. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, *17*(4), 479.

Weaver, J. B., III. (1991). Exploring the links between personality and media preferences. *Personality and Individual Differences*, *12*(12), 1293–1299.

Weaver, J. B., III, & Kiewitz, C. (2007). Eysenck personality questionnaire. In R. A. Reynods, R. Woods, & J. D. Baker (Eds.), *Handbook of Electronic Surveys and Measurements,* (pp. 360–363). Hershey, PA: Idea Group Reference.

Williams, R. L., & Eggert, A. C. (2003). Notetaking in college classes: Student patterns and instructional strategies. *The Journal of General Education*, *51*(3), 173–199.

Yamamoto, K. (2007). Banning laptops in the classroom: Is it worth the hassles? *Journal of Legal Education*, *57*(4), 477–520.

Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit–formation. *Journal of Comparative Neurology and Psychology*, *18*(5), 459–482.

Young, J. R. (2004). When good technology goes bad. *The Chronicle of Higher Education*, (November 12).

Young, J. R. (2006). The fight for classroom attention: Professor vs. laptop. *The Chronicle of Higher Education*, *52*(39), A27.

YouTube. (2010). *Professor Freezes Laptop In Liquid Nitrogen*. Retrieved February 12, 2012 from http://www.youtube.com/watch?v=HyvT_CXMoog

Zucker, A. (2004). Developing a research agenda for ubiquitous computing in schools. *Journal of Educational Computing Research*, *30*(4), 371–386.

Zuckerberg, M. (2010, September 15). *500 Million Stories*. Message posted to http://blog.facebook.com/blog.php?post=409753352130.